

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 22, 2003, 12:17:51 ; Search time 2751 Seconds

(without alignments)
6379.132 Million cell updates/sec

Title: US-09-735-712-1

Perfect score: 603
Sequence: 1 atggaattcaagacacccgcaca.....atttgacaatgttgttga 603

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:
1: gb_ba:*
2: gb_hig:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_ph:*
7: gb_pl:*
8: gb_pl:*
9: gb_pl:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_ju:*
20: em_cm:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hig_hum:*
31: em_hig_inv:*
32: em_hig_other:*
33: em_hig_mus:*
34: em_hig_pln:*
35: em_hig_rod:*
36: em_hig_mam:*
37: em_hig_vrt:*
38: em_sy:*
39: em_higo_hum:*
40: em_higo_mus:*
41: em_higo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	603	100.0	603	6	AX179781	AX179781 Sequence
2	603	100.0	676	6	AX179789	AX179789 Sequence
3	599.8	99.5	762	9	BC029884	BC029884 Homo sapi
4	598.2	99.2	691	9	AB013103	AB013103 Homo sapi
5	598.2	99.2	694	9	AF237907	AF237907 Homo sapi
6	598.2	99.2	760	6	AX268522	AX268522 Sequence
7	596.6	98.9	697	9	AF321127	AF321127 Homo sapi
8	597.4	57.6	450	6	AX179787	AX179787 Sequence
9	341.8	56.7	417	6	AX179785	AX179785 Sequence
10	340.2	56.4	382	6	AX335001	AX335001 Sequence
11	246	40.8	246	6	AX179783	AX179783 Sequence
12	157	26.0	138097	2	AC027787	AC027787 Homo sapi
13	157	26.0	147788	2	AC015840	AC015840 Homo sapi
14	157	26.0	161238	9	AP003127	AP003127 Homo sapi
15	157	26.0	166793	2	AC090401	AC090401 Homo sapi
16	157	25.4	167934	9	AP001034	AP001034 Homo sapi
17	153.4	25.4	138097	2	AC027787	AC027787 Homo sapi
18	106.8	17.7	107211	2	AC111330	AC111330 Rattus no
19	88	14.6	1076	9	AF068288	AF068288 Homo sapi
20	86.4	14.3	720	9	AB022821	AB022821 Homo sapi
21	86.4	14.3	916	9	AB013102	AB013102 Homo sapi
22	86.4	14.3	1550	9	BC020648	BC020648 Homo sapi
23	86.4	14.3	1619	9	AF237912	AF237912 Homo sapi
24	86.4	14.3	1669	6	AR035695	AR035695 Sequence
25	86.4	14.3	1669	6	AR083580	AR083580 Sequence
26	83.2	13.8	959	9	AF350500	AF350500 Homo sapi
27	77.8	12.9	107211	2	AC111330	AC111330 Rattus no
28	76.2	12.6	2805	9	AF367473	AF367473 Homo sapi
29	63.2	10.5	1646	9	HUMIERB	L35848 Homo sapien
30	63.2	10.5	1661	6	A68627	A68627 Sequence 3
31	63.2	10.5	1661	6	AR082014	AR082014 Sequence
32	63.2	10.5	1661	6	BD003343	BD003343 HT m4 , m
33	63.2	10.5	1661	6	I25809	I25809 Sequence 3
34	62.6	10.4	522	9	AF354928	AF354928 Homo sapi
35	62.4	10.3	1222	4	SSR236932	AJ236932 Sus scrofa
36	61.6	10.2	1628	9	BC008487	BC008487 Homo sapi
37	58.8	9.8	2997	9	AK057418	AK057418 Homo sapi
38	58.6	9.7	708	6	AX101306	AX101306 Sequence
39	58.6	9.7	1010	10	MUSFCERB	J05019 Mouse mast
40	56.8	9.4	950	10	AF321128	AF321128 Mus muscu
41	55.6	9.2	899	10	AF237910	AF237910 Mus muscu
42	55.4	9.2	1163	10	AF237915	AF237915 Mus muscu
43	55	9.1	1474	6	I07272	I07272 Sequence 19
44	55	9.1	1474	6	HSCD20	X12530 Human mRNA
45	55	9.1	1476	6	I07277	I07277 Sequence 24

ALIGNMENTS

RESULT 1	AX179781	603 bp	DNA	linear	PAT 06-AUG-2001
LOCUS	AX179781				
DEFINITION	Sequence 1 from Patent WO0146417.				
ACCESSION	AX179781				
VERSION	AX179781.1	GI:15132144			
KEYWORDS					
SOURCE					
ORGANISM	human.				
REFERENCE	Homo sapiens				
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
TITLE	Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.				
	1 (bases 1 to 603)				
	Walke,D.W. and Turner,C.A.				
	Human membrane proteins and polynucleotides encoding the same				
	having homology to cd20 proteins and ige receptors				

```

JOURNAL
Patient: WO 0146117 A 1 28-JUN-2001;
Lexicon Genetics Incorporated (US)
Location/Qualifiers
1. .603
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT      162 a      104 g      212 t
ORIGIN

```

Query Match	100.0%;	Score 603;	DB 6;	Length 603;
Best Local Similarity	100.0%;	Pred. No. 6.2e-129;		
Matches 603; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

QY	61	TCGAATATGAGTCACAGAACTTTTCAGCCAGACGACCTTTTCACTCAAAAGCCCTTGCAA	12
Db	61	TCGAATATGAGTCACAGAACTTTTCAGCCAGACGACCTTTTCAACTCAAAAGCCCTTGCAA	12
QY	121	AAATTAATTTGGCTGGAATAAATCTTAGGGACATACGAGATCCGTGTTGGAATTATG	18
Db	121	AAATTAATTTGGCTGGAATAAATCTTAGGGACATACGAGATCCGTGTTGGAATTATG	18
QY	181	ACCTTTTCCTTTGGAGTATCTTCCTTTTCACTTGTTTAAAAACCATATCCAGAGTTTCCC	24
Db	181	ACCTTTTCCTTTGGAGTATCTTCCTTTTCACTTGTTTAAAAACCATATCCAGAGTTTCCC	24
QY	241	TTTATATTTCTTTCAGAGATATCCATATTCGAGGAGCTCTGTGTTGTTCAATTAATTCGAGGC	30
Db	241	TTTATATTTCTTTCAGAGATATCCATATTCGAGGAGCTCTGTGTTGTTCAATTAATTCGAGGC	30
QY	301	TTCTTAATGCGATGAAAGAAAACACACAAACTCTGTAATATGAGCCGAATAAG	36
Db	301	TTCTTAATGCGATGAAAGAAAACACACAAACTCTGTAATATGAGCCGAATAAG	36
QY	361	AATCTTCTTAGTGCCCTGAGAGCAATACCTGGAATCAATCTCTCAGATTTGGTTTATC	42
Db	361	AATCTTCTTAGTGCCCTGAGAGCAATACCTGGAATCAATCTCTCAGATTTGGTTTATC	42
QY	421	CTGATGATCAAAACTACATTTGTGTGTTATTTCTACCAAAATATGCTAGTGTAAAGCTGTACT	48
Db	421	CTGATGATCAAAACTACATTTGTGTGTTATTTCTACCAAAATATGCTAGTGTAAAGCTGTACT	48
QY	481	GTCCTGTCTTGGGAATTTTGATATCAATGATGACTTTCAGCAATATATGAAATTAATCATT	54
Db	481	GTCCTGTCTTGGGAATTTTGATATCAATGATGACTTTCAGCAATATATGAAATTAATCATT	54
QY	541	TCCTCTCCCTTTCTCAATTTTGGGGTGCCACATCAGAGAAATGTGATGTGAAACAATGTGT	60
Db	541	TCCTCTCCCTTTCTCAATTTTGGGGTGCCACATCAGAGAAATGTGATGTGAAACAATGTGT	60
QY	601	TGA 603	
Db	601	TGA 603	

RESULT 2				
AX179789				
LOCUS	AX179789	676 bp	DNA	linear
DEFINITION	Sequence 9 from Patent WO0146417.			
ACCESSION	AX179789			
VERSION	AX179789.1	GI:15132148		
KEYWORDS				
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiinae; Homo.			
TITLE	1 (bases 1 to 676)			
	Walke,D.W. and Turner,C.A.			
	Human membrane proteins and polynucleotides encoding the same			
	having homology to cd20 proteins and 1ge receptors			

JOURNAL				
PATENT:	WO 0146417 A 9	28-JUN-2001		
LEXICON	Genetics Incorporated (US)			
LOCATION/QUALIFIERS				
1..676				
/organism="Homo sapiens"				
/db_xref="taxon:9606"				
BASE COUNT	193	a	140	c 112 g 231
ORIGIN				

Query Match	100.0%	Score 603;	DB 6;	Length 676;
Best Local Similarity	100.0%	Pred. No. 6.2e-129;		
Matches 603; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

QY	1	ATGGATTCAACACACCGCCACACAGTCCGGTGTTCGTGATTTCTCCAGAAATACAGCT	60
Db	33	ATGCGATTCAACACACCGCCACACAGTCCGGTGTTCGTGATTTCTCCAGAAATACAGCT	92
QY	61	TCAGATATATGAGTCACAGAACTTTCAGCCACGACTTTTCACATCAAGGCCCTTGCAA	120
Db	93	TCAGATATATGAGTCACAGAACTTTCAGCCACGACTTTTCACATCAAGGCCCTTGCAA	152
QY	121	AAATTTATTTGCTAGAAAAATAAAAATCTTAGGGACATCCAGATCCGTGTTGGAAATTATG	180
Db	153	AAATTTATTTGCTAGAAAAATAAAAATCTTAGGGACATCCAGATCCGTGTTGGAAATTATG	212
QY	181	ACCTTTCTCTTTGGAGTATCTTCCTTTTCACCTTGTTAAACCATATCCAAAGSTTCCC	240
Db	213	ACCTTTCTCTTTGGAGTATCTTCCTTTTCACCTTGTTAAACCATATCCAAAGSTTCCC	272
QY	241	TTTATATTTCTTTCAGAGTATCATCTTCGGGGCTCTGTTTGTTCATTAAATTCGAGCC	300
Db	273	TTTATATTTCTTTCAGAGTATCATCTTCGGGGCTCTGTTTGTTCATTAAATTCGAGCC	332
QY	301	TTTCTAATTCGAGAGAAAAAAGCAACCAAGAAATCTGATTAATTTAGACCCGAATATG	360
Db	333	TTTCTAATTCGAGAGAAAAAAGCAACCAAGAAATCTGATTAATTTAGACCCGAATATG	392
QY	361	AATCTTCTTAGTGCCTCGAGAGCAATATGTCGGAATCATATTCCTCCACATTTGGTTATAC	420
Db	393	AATCTTCTTAGTGCCTCGAGAGCAATATGTCGGAATCATATTCCTCCACATTTGGTTATAC	452
QY	421	CTAGATCAAAACTCATATTTGTGTATTTCTCACCAAAATAGTCAGTGTAGGCTGTACT	480
Db	453	CTAGATCAAAACTCATATTTGTGTATTTCTCACCAAAATAGTCAGTGTAGGCTGTACT	512
QY	481	GTCGCTGCTTGGGAATTTTATTAACAATGATGACTTCCTGACATTAATTAATTAAT	540
Db	513	GTCGCTGCTTGGGAATTTTATTAACAATGATGACTTCCTGACATTAATTAATTAAT	572
QY	541	TCTCTGCCCTTCTCAATTTTGGGTGCACACTCAGAGGATTTGATTTGGAACATGTTGT	600
Db	573	TCTCTGCCCTTCTCAATTTTGGGTGCACACTCAGAGGATTTGATTTGGAACATGTTGT	632
QY	601	TGA 603	
Db	633	TGA 635	

RESULT	3
LOCUS	BC029884
DEFINITION	762 bp mRNA linear PRI 20-MAY-2002
ACCESSION	Homo sapiens, membrane-spanning 4-domains, subclass A, member 5, clone MGC:34184 IMAGE:5171526, mRNA, complete cds.
VERSION	BC029884 BC029884.1 GI:20988638
KEYWORDS	MGC.
SOURCE	Homo sapiens. Homo sapiens
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 762)
AUTHORS	Strausberg,R
TITLE	Direct Submission

JOURNAL

Submitted (06-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amgbc@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulvik, S.W., Hale, S.M., Toom, V.S., Kowls, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAC Plate: 51 Row: F Column: 13
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 12965204.

FEATURES

source

1..762
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="MGC:34184 IMAGE:5171526"
 /tissue_type="Brain, adult medulla"
 /clone_1ib="NIH_MGC_119"
 /lab_host="DH10B"
 /note="Vector: PCMV-SPORT6"
 87..689
 /product="membrane-spanning 4-domains, subfamily A, member 5"
 /protein_id="AAH29884.1"
 /db_xref="GI:20988639"
 /translation="MDSSTAHSPVFLVPPETASVESESTELSAFTTSPLOKIFA
 RKMKILGTIOILFGIMTFSFYIFLTLKPYRPPPIFLSGYPMGSVLFINSGLF
 IAVKRRTETLILISRLNPLSALGAIIGIILTFGLIDONYICGSHONSQCAVAV
 VFLGILITLMFSLIELFISLPSFSLGSHSDCCDEQCC"

CDS

BASE COUNT 247 a 152 c 124 g 239 t
 ORIGIN

Query Match 99.5%; Score 599.8; DB 9; Length 762;
 Best Local Similarity 99.7%; Pred. No. 3.3e-128;
 Matches 601; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 ATGGATTCAAGCAGCCACAGTCCGGTGTTCGTGATTTCTCCAGAAATCACTGCT 60
 |||||||
 87 ATGGATTCAAGCAGCCACAGTCCGGTGTTCGTGATTTCTCCAGAAATCACTGCT 146
 |||||||
 61 TCAGAAATAGAGTCACAGAACTTTCAAGCAGACCTTTTCACTCAAAAGCCCTTGCAA 120
 |||||||
 147 TCAGAAATAGAGTCACAGAACTTTCAAGCAGACCTTTTCACTCAAAAGCCCTTGCAA 206
 |||||||
 121 AAATTTATTTGCTAGAAAATGAATCTTAGGAGTATCCAGATCCGTTTGGAAATTAG 180
 |||||||
 207 AAATTTATTTGCTAGAAAATGAATCTTAGGAGTATCCAGATCCGTTTGGAAATTAG 266
 |||||||
 181 ACCTTTCTTTTGGAGTATCTTCTTTCACCTGTAAACATATCCAAAGTTTCCC 240
 |||||||
 267 ACCTTTCTTTTGGAGTATCTTCTTTCACCTGTAAACATATCCAAAGTTTCCC 326
 |||||||
 241 TTATATATTTCTTACAGATATCATCTCTGGGCTCTGTTTGTTCATTAATTTGAGGCC 300
 |||||||
 327 TTATATATTTCTTACAGATATCATCTCTGGGCTCTGTTTGTTCATTAATTTGAGGCC 386
 |||||||
 301 TTCTTAATTCAGTGAAGAAAACACAGAAATCTGTATATTTGAGCCGATTAATG 360
 |||||||

Db 387 TTCTAATTCAGTGAAGAAAACACAGAAATCTGTATATTTAGCCGATTAATG 446
 Oy 361 AATCTTCTTAGTGGCCCTGAGACGAATAGCTGGAATCATTTCTCTACATTTGGTTTCAATC 420
 |||
 Db 447 AATTTCTTAGTGGCCCTGAGACGAATAGCTGGAATCATTTCTCTACATTTGGTTTCAATC 506
 Oy 421 CTGATCAAAACATACATTTGTTGTTATTTCTACCAAAATATGTCAGTATAGGCTGTACT 480
 |||||||
 Db 507 CTGATCAAAACATACATTTGTTGTTATTTCTACCAAAATATGTCAGTATAGGCTGTACT 566
 Oy 481 GTCTGTTCTTGGGAATTTTGAATTAATGATGATCTTTCAGATTTTGAATTAATTCAT 540
 |||||||
 Db 567 GTCTGTTCTTGGGAATTTTGAATTAATGATGATCTTTCAGATTTTGAATTAATTCAT 626
 Oy 541 TCTCTGCTTTTCAATTTTGGGCTGCTACAGAGATTTGATTTGTAACAATTTGT 600
 |||||||
 Db 627 TCTCTGCTTTTCAATTTTGGGCTGCTACAGAGATTTGATTTGTAACAATTTGT 686

RESULT 4

AB013103 691 bp mRNA linear PRI 20-MAR-2001

LOCUS AB013103 Homo sapiens mRNA for MS4A5, complete cds.
 DEFINITION
 AB013103.1 GI:11559213
 ACCESSION
 MS4A5: CD20-like 2.
 VERSION
 MS4A5: CD20-like 2.
 KEYWORDS
 Homo sapiens testis cDNA to mRNA.
 SOURCE
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (sites)
 Ishibashi, K., Suzuki, M., Sasaki, S. and Imai, M.
 Identification of a new multigene four-transmembrane family (MS4A)
 related to CD20, Htm4 and beta subunit of the high-affinity Ige
 receptor
 JOURNAL
 Gene 264 (1), 87-93 (2001)
 MEDLINE
 2 (sites)
 REFERENCE
 Ishibashi, K., Sasaki, S. and Marumo, F.
 Cloning of three CD20 homolog from human, putative calcium channels
 JOURNAL
 Unpublished
 3 (bases 1 to 691)
 REFERENCE
 Ishibashi, K.
 Direct Submission
 Submitted (20-APR-1998) Kenichi Ishibashi, Tokyo Medical and Dental
 University, 2nd Internal Medicine; yushima 1-5-45, Bunkyo, Tokyo
 113-8519, Japan (E-mail: kishibashi.med2@med.tmd.ac.jp,
 Tel:81-3-5803-5223, Fax:81-3-5803-0132)
 FEATURES
 source
 1..691
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /tissue_type="testis"
 1..691
 /gene="MS4A5"
 52..654
 /gene="MS4A5"
 /codon_start=1
 /product="MS4A5"
 /protein_id="BAB18739.1"
 /db_xref="GI:11559214"
 /translation="MDSSTAHSPVFLVPPETASVESESTELSAFTTSPLOKIFA
 RKMKILGTIOILFGIMTFSFYIFLTLKPYRPPPIFLSGYPMGSVLFINSGLF
 IAVKRRTETLILISRLNPLSALGAIIGIILTFGLIDONYICGSHONSQCAVAV
 VFLGILITLMFSLIELFISLPSFSLGSHSDCCDEQCC"

BASE COUNT 195 a 142 c 117 g 237 t
 ORIGIN
 Query Match 99.2%; Score 598.2; DB 9; Length 691;

Best Local Similarity 99.5%; Pred. No. 7.9e-128;
Matches 600; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
QY 1 ATGATTTCAAGCACCACAGCAGTCCGGTGTCTGTGATTTCTCCAGAAATCAGTGT 60
    |||
DB 52 ATGATTTCAAGCACCACAGCAGTCCGGTGTCTGTGATTTCTCCAGAAATCAGTGT 111
    |||
QY 61 TCAGAAATAGAGTCACAGAACTTTCAGCAGACCTTTTCAACGCAAAAGCCCTTGCAA 120
    |||
DB 112 TCAGAAATAGAGTCACAGAACTTTCAGCAGACCTTTTCAACGCAAAAGCCCTTGCAA 171
    |||
QY 121 AAATTTATTTCTAGAAAATGAAAATCTTAGGACATATCCAGATCCTGTTGGAATTATG 180
    |||
DB 172 AAATTTATTTCTAGAAAATGAAAATCTTAGGACATATCCAGATCCTGTTGGAATTATG 231
    |||
QY 181 ACCTTTCTTTTGAGATATCTTCCTTTTCACTTTTCACTTTTCACTTTTCACTTTTCC 240
    |||
DB 232 ACCTTTCTTTTGAGATATCTTCCTTTTCACTTTTCACTTTTCACTTTTCACTTTTCC 291
    |||
QY 241 TTTATATTTCTTTCAGATATCCATTCCTGAGGCTCTGTTTGTTCATTAATTCGAGACC 300
    |||
DB 292 TTTATATTTCTTTCAGATATCCATTCCTGAGGCTCTGTTTGTTCATTAATTCGAGACC 351
    |||
QY 301 TTCTTAATTTGCACTGAAAAGAAAACACAGAACTCTGTAAATATTGAGCCGAATATAG 360
    |||
DB 352 TTCTTAATTTGCACTGAAAAGAAAACACAGAACTCTGTAAATATTGAGCCGAATATAG 411
    |||
QY 361 AATCTTCTTTTGAGATATCTTCCTTTTCACTTTTCACTTTTCACTTTTCACTTTTCC 420
    |||
DB 412 AATTTTCTTTTGAGATATCTTCCTTTTCACTTTTCACTTTTCACTTTTCACTTTTCC 471
    |||
QY 421 CTAGATCAAAACTACATTTGTGTATTTCTACACAAAATAGTCAAGTAAAGCTGTACT 480
    |||
DB 472 CTAGATCAAAACTACATTTGTGTATTTCTACACAAAATAGTCAAGTAAAGCTGTACT 531
    |||
QY 481 GTCTGTCTTCTGGGAATTTTGATTAATGATGATCACTTTCAGCAATTATGCAATTTTCAAT 540
    |||
DB 532 GTCTGTCTTCTGGGAATTTTGATTAATGATGATCACTTTCAGCAATTATGCAATTTTCAAT 591
    |||
QY 541 TCTCTGCTTTTCAATTTTGGGGTGCACATCAGAGATGTGATTTGTAACATGTGT 600
    |||
DB 592 TCTCTGCTTTTCAATTTTGGGGTGCACATCAGAGATGTGATTTGTAACATGTGT 651
    |||
QY 601 TGA 603
    |||
DB 652 TGA 654
    |||
RESULT 5
AF237907 694 bp mRNA linear PRI 17-APR-2001
LOCUS Homo sapiens MS4A5 protein mRNA, complete cds.
DEFINITION AF237907
ACCESSION AF237907
VERSION AF237907.1 GI:13649400
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 694)
AUTHORS Liang, Y. and Tedder, T.F.
TITLE Identification of a CD20-, FcepsilonR1beta-, and Htm4-related gene
family: sixteen new MS4A family members expressed in human and
mouse
JOURNAL Genomics 72 (2), 119-127 (2001)
MEDLINE 21295030
PUBMED 11401424
REFERENCE 2 (bases 1 to 694)
AUTHORS Liang, Y. and Tedder, T.F.
TITLE Direct Submission
JOURNAL Submitted (24-FEB-2000) Immunology, Duke Medical Center, Research
Dr., Durham, NC 27710, USA
FEATURES
Location/Qualifiers
```

source

1..694
/organism="Homo sapiens"
/db_xref="taxon:9606"
55..657
/codon_start=1
/product="MS4A5 protein"
/protein_id="AAK37416.1"
/db_xref="GI:13649401"

CDS

/translation="MDSYAHSPVFLVPEETASEVESTELSATFSPLOKIFA
RKMLIGTIOLEFGIMFSEGVFLFTLLKPYPRPFLISGVPMGSVFLNSGAF
LAVKRTETLILSRIMNPLSALGAILILHFGFLIDONYIGSYHNSQCAVY
VLEFLILITMFSTIELEFISPSIILGSHSDCCBQCC"

BASE COUNT 198 a 143 c 116 g 237 t
ORIGIN

Query Match 99.2%; Score 598.2; DB 9; Length 694;
Best Local Similarity 99.5%; Pred. No. 7.9e-128;

Matches 600; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
QY 1 ATGATTTCAAGCACCACAGCAGTCCGGTGTCTGTGATTTCTCCAGAAATCAGTGT 60
    |||
DB 55 ATGATTTCAAGCACCACAGCAGTCCGGTGTCTGTGATTTCTCCAGAAATCAGTGT 114
    |||
QY 61 TCAGAAATAGAGTCACAGAACTTTCAGCAGACCTTTTCAACGCAAAAGCCCTTGCAA 120
    |||
DB 115 TCAGAAATAGAGTCACAGAACTTTCAGCAGACCTTTTCAACGCAAAAGCCCTTGCAA 174
    |||
QY 121 AAATTTATTTCTAGAAAATGAAAATCTTAGGACATATCCAGATCCTGTTGGAATTATG 180
    |||
DB 175 AAATTTATTTCTAGAAAATGAAAATCTTAGGACATATCCAGATCCTGTTGGAATTATG 234
    |||
QY 181 ACCTTTCTTTTGAGATATCTTCCTTTTCACTTTTCACTTTTCACTTTTCACTTTTCC 240
    |||
DB 235 ACCTTTCTTTTGAGATATCTTCCTTTTCACTTTTCACTTTTCACTTTTCACTTTTCC 294
    |||
QY 241 TTTATATTTCTTTCAGATATCCATTCCTGAGGCTCTGTTTGTTCATTAATTCGAGACC 300
    |||
DB 295 TTTATATTTCTTTCAGATATCCATTCCTGAGGCTCTGTTTGTTCATTAATTCGAGACC 354
    |||
QY 301 TTCTTAATTTGCACTGAAAAGAAAACACAGAACTCTGTAAATATTGAGCCGAATATAG 360
    |||
DB 355 TTCTTAATTTGCACTGAAAAGAAAACACAGAACTCTGTAAATATTGAGCCGAATATAG 414
    |||
QY 361 AATCTTCTTTTGAGATATCTTCCTTTTCACTTTTCACTTTTCACTTTTCACTTTTCC 420
    |||
DB 415 AATTTTCTTTTGAGATATCTTCCTTTTCACTTTTCACTTTTCACTTTTCACTTTTCC 474
    |||
QY 421 CTAGATCAAAACTACATTTGTGTATTTCTACACAAAATAGTCAAGTAAAGCTGTACT 480
    |||
DB 475 CTAGATCAAAACTACATTTGTGTATTTCTACACAAAATAGTCAAGTAAAGCTGTACT 534
    |||
QY 481 GTCTGTCTTCTGGGAATTTTGATTAATGATGATCACTTTCAGCAATTATGCAATTTTCAAT 540
    |||
DB 535 GTCTGTCTTCTGGGAATTTTGATTAATGATGATCACTTTCAGCAATTATGCAATTTTCAAT 594
    |||
QY 541 TCTCTGCTTTTCAATTTTGGGGTGCACATCAGAGATGTGATTTGTAACATGTGT 600
    |||
DB 595 TCTCTGCTTTTCAATTTTGGGGTGCACATCAGAGATGTGATTTGTAACATGTGT 654
    |||
QY 601 TGA 603
    |||
DB 655 TGA 657
    |||
RESULT 6
AX268522 760 bp DNA linear PAT 29-OCT-2001
LOCUS AX268522
DEFINITION Sequence 1 from Patent WO0174903.
ACCESSION AX268522
VERSION AX268522.1 GI:16541676
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
```


Db 357 TTCTAAATTCAGTGAAGAAAAAGAACACAGAACTGTGATATATGAGCCGAATTAATG 416
QY 361 AATCTTCTAGTCCCTGAGCAATAGCTGAATCATTCCTGACATTTGGTTTCATC 420
Db 417 AATTTCTTAGTCCCTGGAGCAATAGCTGAATCATTCCTGACATTTGGTTTCATC 476
QY 421 CTAGATCAAAACTACATTTGTTGTTATTTCTACCAAAATAGTCACTGTAAGGCTGTACT 480
Db 477 CTAGATCAAAACTACATTTGTTGTTATTTCTACCAAAATAGTCACTGTAAGGCTGTACT 536
QY 481 GTCCGTCTCTGGGAATTTTGATTTACATTTGATGACTTTTCAGATTAATTAATTCATT 540
Db 537 GTCCGTCTCTGGGAATTTTGATTTACATTTGATGACTTTTCAGATTAATTAATTCATT 596
QY 541 TCTGTCCCTTCTCAATTTTGGGTGCCACAGAGATTTGATTTGTAACAAATGTTGT 600
Db 597 TCTGTCCCTTCTCAATTTTGGGTGCCACAGAGATTTGATTTGTAACAAATGTTGT 656
QY 601 TGA 603
Db 657 TGA 659

RESULT 8
AX179787 450 bp DNA linear PAT 06-AUG-2001
LOCUS Sequence 7 from Patent WO0146417.
DEFINITION AX179787
ACCESSION AX179787
VERSION AX179787.1 GI:15132147
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 450)
AUTHORS Walke,D.W. and Turner,C.A.
TITLE Human membrane proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0146417-A 7 28 -JUN-2001;
Lexicon Genetics Incorporated (US)
FEATURES
Source location/Qualifiers
1..450
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 120 a 94 c 77 g 159 t
ORIGIN

Query Match 57.6%; Score 347.4; DB 6; Length 450;
Best Local Similarity 97.0%; Pred. No. 5.6e-70;
Matches 354; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 ATGATTCAGACCGCACACAGTCCGGTGTTCGTATTTCCCTCCAGAAATCACTGCT 60
Db 1 ATGATTCAGACCGCACACAGTCCGGTGTTCGTATTTCCCTCCAGAAATCACTGCT 60
QY 61 TCAGATATGAGTCCACAGAACTTTACGACAGACCTTTTCACTCAAAAGCCCTTGCA 120
Db 61 TCAGATATGAGTCCACAGAACTTTACGACAGACCTTTTCACTCAAAAGCCCTTGCA 120
QY 121 AAATATTTTGTAGAAAAATGAAATCTTAGGACTATCCAGATCCTGTTTGAATTAATG 180
Db 121 AAATATTTTGTAGAAAAATGAAATCTTAGGACTATCCAGATCCTGTTTGAATTAATG 180
QY 181 ACCTTTCTTTTGGAGTATCTTCTTTACCTGTTTAAACCATATCCAAAGGTTTCC 240
Db 181 ACCTTTCTTTTGGAGTATCTTCTTTACCTGTTTAAACCATATCCAAAGGTTTCC 240
QY 241 TTTATATTTCTTTCAGGATATCCATTTCTGGGCTCTGTTTGTATTAATTCGAGCC 300
Db 241 TTTATATTTCTTTCAGGATATCCATTTCTGGGCTCTGTTTGTATTAATTCGAGCC 300
QY 301 TTCCAAATTCGATGAAAAAGAAAAACACAGAACTGTGATTAATG 360
Db 301 TTCCAAATTCGATGAAAAAGAAAAACACAGAACTGTGATTAATG 360

Db 301 TTCTAAATTCAGTGAAGAAAAAGAACACAGAACTGTGGAATTTGATTAATTCATTGATG 360
QY 361 AATCT 365
Db 361 ACTTT 365

RESULT 9
AX179785 417 bp DNA linear PAT 06-AUG-2001
LOCUS Sequence 5 from Patent WO0146417.
DEFINITION AX179785
ACCESSION AX179785
VERSION AX179785.1 GI:15132146
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 417)
AUTHORS Walke,D.W. and Turner,C.A.
TITLE Human membrane proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0146417-A 5 28 -JUN-2001;
Lexicon Genetics Incorporated (US)
FEATURES
Source location/Qualifiers
1..417
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 115 a 91 c 69 g 142 t
ORIGIN

Query Match 56.7%; Score 341.8; DB 6; Length 417;
Best Local Similarity 99.4%; Pred. No. 1.1e-68;
Matches 343; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGATTCAGACCGCACACAGTCCGGTGTTCGTATTTCCCTCCAGAAATCACTGCT 60
Db 1 ATGATTCAGACCGCACACAGTCCGGTGTTCGTATTTCCCTCCAGAAATCACTGCT 60
QY 61 TCAGATATGAGTCCACAGAACTTTACGACAGACCTTTTCACTCAAAAGCCCTTGCA 120
Db 61 TCAGATATGAGTCCACAGAACTTTACGACAGACCTTTTCACTCAAAAGCCCTTGCA 120
QY 121 AAATATTTTGTAGAAAAATGAAATCTTAGGACTATCCAGATCCTGTTTGAATTAATG 180
Db 121 AAATATTTTGTAGAAAAATGAAATCTTAGGACTATCCAGATCCTGTTTGAATTAATG 180
QY 181 ACCTTTCTTTTGGAGTATCTTCTTTACCTGTTTAAACCATATCCAAAGGTTTCC 240
Db 181 ACCTTTCTTTTGGAGTATCTTCTTTACCTGTTTAAACCATATCCAAAGGTTTCC 240
QY 241 TTTATATTTCTTTCAGGATATCCATTTCTGGGCTCTGTTTGTATTAATTCGAGCC 300
Db 241 TTTATATTTCTTTCAGGATATCCATTTCTGGGCTCTGTTTGTATTAATTCGAGCC 300
QY 301 TTCCAAATTCGATGAAAAAGAAAAACACAGAACTGTGATTAATG 345
Db 301 TTCCAAATTCGATGAAAAAGAAAAACACAGAACTGTGATTAATG 345

RESULT 10
AX335001/c 382 bp DNA linear PAT 09-JAN-2002
LOCUS Sequence 5510 from Patent WO0194629.
DEFINITION AX335001
ACCESSION AX335001
VERSION AX335001.1 GI:18125720
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,

Horigan, S., Soppel, D.R. and Weaver, Z.
 Cancer gene determination and therapeutic screening using signature
 gene sets
 Patent: WO 0194629-A 5510 13-DEC-2001;
 JOURNAL Avalon Pharmaceuticals (US)
 FEATURES 1.382
 Location/Qualifiers
 source /db_xref="taxon:9606"
 BASE COUNT 135 a 73 c 67 g 107 t
 ORIGIN

Query Match 56.4%; Score 340.2; DB 6; Length 382;
 Best Local Similarity 99.1%; Pred. No. 2.6e-68;
 Matches 342; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 259 TATTCATTTCTGGGGCTCTGTTTGTTCATTAATTTGAGACCTTCCTTAATTTGAGTGA 318
 |||||||
 Db 382 TATTCATTTCTGGGGCTCTGTTTGTTCATTAATTTGAGACCTTCCTTAATTTGAGTGA 323
 |||||||

QY 319 AGAAAAACACAGAACTCTGATATATTTGAGCCGAATTAATGAATCTTTAGTCCCTG 378
 |||||||
 Db 322 AGAAAAACACAGAACTCTGATATATTTGAGCCGAATTAATGAATCTTTAGTCCCTG 263
 |||||||

QY 379 AGAGCAATAGCTGGAATCATTTCTCCACATTTGGTTTCATTCCTAGATCAAACTACAT 438
 |||||||
 Db 262 GGAGCAATAGCTGGAATCATTTCTCCACATTTGGTTTCATTCCTAGATCAAACTACAT 203
 |||||||

QY 439 TGTGTTATTTCTCACCAGAAATAGTCAAGCTTCTACTGTCCTGTTCTTGGGAAT 498
 |||||||
 Db 202 TGTGTTATTTCTCACCAGAAATAGTCAAGCTTCTACTGTCCTGTTCTTGGGAAT 143
 |||||||

QY 499 TTGATTACATTTGATGACTTTTCAGCATTAATTAATTAATTTCTCTGCTTCTCAAT 558
 |||||||
 Db 142 TTGATTACATTTGATGACTTTTCAGCATTAATTAATTAATTTCTCTGCTTCTCAAT 83
 |||||||

QY 559 TTGGGGTGCCACACAGAGATTTGATTTGAACAATGTTTGA 603
 |||||||
 Db 82 TTGGGGTGCCACACAGAGATTTGATTTGAACAATGTTTGA 38
 |||||||

RESULT 11
 AX179783 246 bp DNA linear PAT 06-AUG-2001
 LOCUS AX179783
 DEFINITION Sequence 3 from Patent WO0146417.
 ACCESSION AX179783
 VERSION AX179783.1 GI:15132145
 KEYWORDS human.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 246)
 AUTHORS Walke, D.W. and Turner, C.A.
 TITLE Human membrane proteins and polynucleotides encoding the same
 JOURNAL Patent: WO 0146417-A 3 28-JUN-2001;
 FEATURES Lexicon Genetics Incorporated (US)
 source Location/Qualifiers
 1.246
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 60 a 45 c 47 g 94 t
 ORIGIN

Query Match 40.8%; Score 246; DB 6; Length 246;
 Best Local Similarity 100.0%; Pred. No. 1.4e-46;
 Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 358 ATGAATCTTTAGTGGCCCTGAGAGCAATAGCTGGAATCATTTCTCTACATTTGGTTTC 417
 |||||||
 Db 1 ATGAATCTTTAGTGGCCCTGAGAGCAATAGCTGGAATCATTTCTCTACATTTGGTTTC 60
 |||||||

QY 418 ATCTAGATCAAACTACATTTGTGTTATTTCTACCAAAATAGTCAAGTCTGT 477
 |||||||
 Db 61 ATCTAGATCAAACTACATTTGTGTTATTTCTACCAAAATAGTCAAGTCTGT 120
 |||||||

QY 478 ACTGCTGTTCTTGGGAATTTTATATACATTTGATGATTTGACATTAATTAATTTTC 537
 |||||||
 Db 121 ACTGCTGTTCTTGGGAATTTTATATACATTTGATGATTTGACATTAATTAATTTTC 180
 |||||||

QY 538 ATTTCTGCTCTTCTCAATTTTGGGGTCCACTCAGAGATTTGATTTGAACAATGT 597
 |||||||
 Db 181 ATTTCTGCTCTTCTCAATTTTGGGGTCCACTCAGAGATTTGATTTGAACAATGT 240
 |||||||

QY 598 TGTGGA 603
 |||||||
 Db 241 TGTGGA 246
 |||||||

RESULT 12
 AC027787/c 138097 bp DNA linear HTG 24-AUG-2002
 LOCUS AC027787/c
 DEFINITION Homo sapiens chromosome 11 clone RP11-196E16 map 11, LOW-PASS
 SEQUENCE SAMPLING.
 ACCESSION AC027787
 VERSION AC027787.2 GI:9845160
 KEYWORDS HTG; HTGS_PHASE0.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 138097)
 AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
 JOURNAL Homo sapiens chromosome 11, clone RP11-196E16
 REFERENCE 2 (bases 1 to 138097)
 AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
 Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,
 Boguslavskiy, L., Boukhalter, B., Brown, A., Burkett, G.,
 Campolino, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,
 Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S.,
 Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
 Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
 Grand-Pierre, N., Grant, G., Hagos, B., Heatford, A., Horton, L.,
 Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
 Klein, J., Laroque, K., Lamazares, R., Landers, T., Lehoczy, J.,
 Levine, R., Lieu, C., Liu, G., Locke, K., MacDonald, P., Marquis, N.,
 McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,
 Meldrim, J., Mihova, T., Miranda, C., Mienga, V., Morrow, J.,
 Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
 O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,
 Pisanl, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
 Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
 Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
 Testaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,
 Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
 Young, G., Zainoun, J., Zimmer, A. and Zody, M.
 TITLE Direct Submission
 JOURNAL Submitted (01-APR-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 REFERENCE 3 (bases 1 to 138097)
 AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
 Anderson, S., Barna, N., Bastien, V., Beda, F., Boguslavskiy, L.,
 Boukhalter, B., Brown, A., Burkett, G., Campolino, A., Castle, A.,
 Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,
 Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Ferreira, P.,
 Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M.,
 Graham, L., Grand-Pierre, N., Hagos, B., Heatford, A., Horton, L.,
 Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Laroque, K.,
 Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lieu, C., Liu, G.,
 MacDonald, P., Marquis, N., McCarthy, M., McEwan, P., McKernan, K.,
 McPheeters, R., Meldrim, J., Mihova, T., Miranda, C., Mienga, V.,
 Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T.,
 O'Neill, P., O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K.,
 Pierre, N., Pisanl, C., Pollara, V., Raymond, C., Rieback, M., Riley, R.,

TITLE
JOURNAL
COMMENT

Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P.,
Sounez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,
Tirrell, A., Travers, M., Triglio, J., Vassiliev, H., Viel, R., Vo, A.,
Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J.,
Zimmer, A. and Zody, M.

Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 18, 2000 this sequence version replaced gi:7382629.

All repeats were identified using RepeatMasker:

Smith, A.F.A. & Green, P. (1996:1997)

<http://ftp.genome.washington.edu/RN/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: MIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L9008

Center clone name: 196_E_16

* NOTE: This record contains 163 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

* 1 663: contig of 663 bp in length

* 664 763: gap of 100 bp

* 764 1464: contig of 701 bp in length

* 1465 1564: gap of 100 bp

* 1565 2268: contig of 704 bp in length

* 2269 2368: gap of 100 bp

* 2369 3042: contig of 674 bp in length

* 3043 3142: gap of 100 bp

* 3143 3820: contig of 678 bp in length

* 3821 3920: gap of 100 bp

* 3921 4626: contig of 706 bp in length

* 4627 4726: gap of 100 bp

* 4727 5423: contig of 697 bp in length

* 5424 5523: gap of 100 bp

* 5524 6234: contig of 711 bp in length

* 6235 6334: gap of 100 bp

* 6335 7040: contig of 706 bp in length

* 7041 7140: gap of 100 bp

* 7141 7837: contig of 697 bp in length

* 7838 7937: gap of 100 bp

* 7938 8632: contig of 695 bp in length

* 8633 8732: gap of 100 bp

* 8733 9430: contig of 698 bp in length

* 9431 9530: gap of 100 bp

* 9531 10214: contig of 684 bp in length

* 10215 10314: gap of 100 bp

* 10315 10980: contig of 666 bp in length

* 10981 11080: gap of 100 bp

* 11081 11783: contig of 703 bp in length

* 11784 11883: gap of 100 bp

* 11884 12582: contig of 699 bp in length

* 12583 12682: gap of 100 bp

* 12683 13389: contig of 707 bp in length

* 13390 13489: gap of 100 bp

* 13490 14186: contig of 697 bp in length

* 14187 14286: gap of 100 bp

* 14287 14996: contig of 710 bp in length

* 14997 15096: gap of 100 bp

* 15097 15796: contig of 700 bp in length

* 15797 15896: gap of 100 bp

* 15897 16583: contig of 687 bp in length

* 16584 16683: gap of 100 bp

* 16684 17376: contig of 693 bp in length

* 17377 17476: gap of 100 bp

* 17477 18178: contig of 702 bp in length

* 18179 18278: gap of 100 bp

* 18279 18962: contig of 684 bp in length

* 18963 19062: gap of 100 bp

* 19063 19673: contig of 611 bp in length

* 19674 19773: gap of 100 bp

* 19774 20436: contig of 663 bp in length

* 20437 20536: gap of 100 bp

* 20537 21233: contig of 697 bp in length

* 21234 21333: gap of 100 bp

* 21334 22042: contig of 709 bp in length

* 22043 22142: gap of 100 bp

* 22143 22808: contig of 666 bp in length

* 22809 22908: gap of 100 bp

* 22909 23595: contig of 687 bp in length

* 23596 23695: gap of 100 bp

* 23696 24391: contig of 696 bp in length

* 24392 24491: gap of 100 bp

* 24492 25190: contig of 699 bp in length

* 25191 25290: gap of 100 bp

* 25291 25981: contig of 691 bp in length

* 25982 26081: gap of 100 bp

* 26082 26765: contig of 684 bp in length

* 26766 26865: gap of 100 bp

* 26866 27563: contig of 698 bp in length

* 27564 27663: gap of 100 bp

* 27664 28377: contig of 714 bp in length

* 28378 28477: gap of 100 bp

* 28478 29184: contig of 707 bp in length

* 29185 29284: gap of 100 bp

* 29285 29983: contig of 699 bp in length

* 29984 30083: gap of 100 bp

* 30084 30775: contig of 692 bp in length

* 30776 30875: gap of 100 bp

* 30876 31565: contig of 690 bp in length

* 31566 31665: gap of 100 bp

* 31666 32360: contig of 695 bp in length

* 32361 32460: gap of 100 bp

* 32461 33157: contig of 697 bp in length

* 33158 33257: gap of 100 bp

* 33258 33951: contig of 694 bp in length

* 33952 34051: gap of 100 bp

* 34052 34735: contig of 684 bp in length

* 34736 34835: gap of 100 bp

* 34836 35483: contig of 648 bp in length

* 35484 35583: gap of 100 bp

* 35584 36308: contig of 725 bp in length

* 36309 36408: gap of 100 bp

* 36409 37108: contig of 700 bp in length

* 37109 37208: gap of 100 bp

* 37209 37892: contig of 684 bp in length

* 37893 37992: gap of 100 bp

* 37993 38672: contig of 680 bp in length

* 38673 38772: gap of 100 bp

* 38773 39475: contig of 703 bp in length

* 39476 39575: gap of 100 bp

* 39576 40290: contig of 715 bp in length

* 40291 40390: gap of 100 bp

* 40391 41087: contig of 697 bp in length

* 41088 41187: gap of 100 bp

* 41188 41891: contig of 704 bp in length

* 41892 41991: gap of 100 bp

* 41992 42679: contig of 688 bp in length

* 42680 42779: gap of 100 bp

* 42780 43500: contig of 721 bp in length

* 43501 43600: gap of 100 bp

* 43601 44295: contig of 695 bp in length

* 44296 44395: gap of 100 bp

* 44396 45088: contig of 693 bp in length

```

Query Match          26.0%: Score 157; DB 2; Length 138097;
Best Local Similarity 91.7%: Pred. No. 2.6e-26;
Matches 166; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 336 TCTGATTAATATGAGCCCAATATGAAATCTTCTAGTGCCTGAGAGCAATAGCTGGAAT 395
      |||||||
Db 102667 TCAGATTAATATGAGCCCAATATGAAATCTTCTAGTGCCTGAGAGCAATAGCTGGAAT 102608

QY 396 CATTCCTCCATCATTTGGTTTCATCTCTAGATCAAACTACATTTTGGTTATTCACCA 455
      |||||||
Db 102607 CATTCCTCCATCATTTGGTTTCATCTCTAGATCAAACTACATTTTGGTTATTCACCA 102548

QY 456 AAATAGTCAGTGAAGCGCTGTTACTGCTCTGTTGGGAATTTGATTACATTTGATGAC 515
      |||||||
Db 102547 AAATAGTCAGTGAAGCGCTGTTACTGCTCTGTTGGGAATTTGATTACATTTGATGAC 102488

QY 516 T 516
Db 102487 T 102487

RESULT 13
AC015840 147788 bp DNA linear HTG 27-APR-2000
LOCUS Homo sapiens clone RP11-24D1, WORKING DRAFT SEQUENCE, 5 unordered
DEFINITION pieces:
AC015840
VERSION AC015840.2 GI:7657730
KEYWORDS HTG; PHASE; HTGS; DRAFT.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 147788)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens, clone RP11-24D1
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 147788)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckert,J., Boguslavsky,L., Boukhalter,B.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., DeArrellano,K., Dewar,K., Domino,M., Donegan,L., Doyle,M.,
Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardina,S., Grant,G., Hagos,B., Heathford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lemock,J., Liu,C., Locke,K., Macdonald,P., Marquis,N.,
McMan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Tittell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
DIRECT SUBMISSION
Submitted (17-NOV-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 27, 2000 this sequence version replaced g1:6446828.
All repeats were identified using RepeatMasker:
Smt,A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: W1BR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center Project name: L4434
Center clone name: 24.D.1
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 136347 bases at least Q40

```

```

Consensus quality: 143600 bases at least Q30
Consensus quality: 146212 bases at least Q20
Insert size: 203000; agarose-fp
Insert size: 147388; sum-of-ctnigs
Quality coverage: 5.2 in Q20 bases; agarose-fp
Quality coverage: 7.1 in Q20 bases; sum-of-ctnigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1. 6594: contig of 6594 bp in length
* 6595 6694: gap of 100 bp
* 6695 16187: contig of 9493 bp in length
* 16188 16287: gap of 100 bp
* 16288 25089: contig of 8802 bp in length
* 25090 25189: gap of 100 bp
* 25190 63923: contig of 38734 bp in length
* 63924 64023: gap of 100 bp
* 64024 147788: contig of 83765 bp in length.
* Location/Qualifiers
  1. 147788
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="RP11-24D1"
    /clone_1id="RP11-11 Human Male BAC"
    1. 6594
    /note="assembly-fragment
    clone_end:77
    vector_slide:left"
    misc_feature
    /note="assembly-fragment"
    16288..25089
    /note="assembly-fragment
    clone_end:SP6
    vector_slide:right"
    25190..63923
    /note="assembly-fragment"
    64024..147788
    /note="assembly-fragment"
BASE COUNT 45466 a 29482 c 28861 g 43569 t 410 others
ORIGIN
Query Match          26.0%: Score 157; DB 2; Length 147788;
Best Local Similarity 91.7%: Pred. No. 2.5e-26;
Matches 166; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 336 TCTGATTAATATGAGCCCAATATGAAATCTTCTAGTGCCTGAGAGCAATAGCTGGAAT 395
      |||||||
Db 132133 TCAGATTAATATGAGCCCAATATGAAATCTTCTAGTGCCTGAGAGCAATAGCTGGAAT 132192

QY 396 CATTCCTCCATCATTTGGTTTCATCTCTAGATCAAACTACATTTTGGTTATTCACCA 455
      |||||||
Db 132193 CATTCCTCCATCATTTGGTTTCATCTCTAGATCAAACTACATTTTGGTTATTCACCA 132252

QY 456 AAATAGTCAGTGAAGCGCTGTTACTGCTCTGTTGGGAATTTGATTACATTTGATGAC 515
      |||||||
Db 132253 AAATAGTCAGTGAAGCGCTGTTACTGCTCTGTTGGGAATTTGATTACATTTGATGAC 132212

QY 516 T 516
Db 132313 T 132313

RESULT 14
AP003127 161238 bp DNA linear PRI 20-DEC-2001
LOCUS Homo sapiens genomic DNA, chromosome 11q clone:RP11-71066, complete
DEFINITION sequences.

```

ACCESSION AP003127
 VERSION AP003127.2 GI:17939962
 KEYWORDS HTG.
 SOURCE Homo sapiens DNA, clone:RP11-729B6.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
 AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 TITLE Homo sapiens genomic DNA
 JOURNAL Published only in Database (2001)
 REFERENCE 2 (bases 1 to 161238)
 AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 TITLE Direct Submission
 JOURNAL Submitted (26-JAN-2001) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-Chou,Tsukumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:hattori@gscc.riken.go.jp, URL:http://hgp.gscc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)

COMMENT On Dec 19, 2001 this sequence version replaced gi:12597183.
 FEATURES
 SOURCE Location/Qualifiers
 1..161238
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11q"
 /clone="RP11-729B6"
 BASE COUNT 50087 a 31927 c 31564 g 47660 t
 ORIGIN

Query Match 26.0%; Score 157; DB 9; Length 161238;
 Best Local Similarity 91.7%; Pred. No. 2,5e-26;
 Matches 166; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 336 TCTGATATATTTGAGCGGATATGATCTTTCTTCTGAGAGCAATGCTGGAAT 395
 |||||||
 Db 137682 TCACATATATTTGAGCGGATATGATCTTTCTTCTGAGAGCAATGCTGGAAT 137741
 |||||||

QY 396 CATCTCTCTCATTTGGTTTCATCTCAGATCAAAACATCTTGTGATTTCATCAACA 455
 |||||||
 Db 137742 CATCTCTCTCATTTGGTTTCATCTCAGATCAAAACATCTTGTGATTTCATCAACA 137801
 |||||||

QY 456 AATAGTCAGTGAAGGCTGTTACTGTCCTGTTGGATTTGATTACATGATGAC 515
 |||||||
 Db 137802 AATAGTCAGTGAAGGCTGTTACTGTCCTGTTGGATTTGATTACATGATGAC 515
 |||||||

QY 516 T 516
 Db 137862 T 137862

RESULT 15
 AC090401/c 166793 bp DNA linear HTG 10-APR-2002
 LOCUS Homo sapiens chromosome 11 clone RP11-729B4 map 11, *** SEQUENCING
 DEFINITION IN PROGRESS ***, 1 ordered piece.
 ACCESSION AC090401
 VERSION AC090401.5 GI:20128179
 KEYWORDS HTG; HTGS_PHASE2; HTGS_FULFILL; HTGS_ACTIVEFIN.
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 166793)
 AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
 TITLE Homo sapiens chromosome 11, clone RP11-729B4
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 166793)
 AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S., Barnes,N., Bastien,V., Boguslavsky,L., Boukhalter,B., Brown,A., Camarata,J., Choepel,Y., Colangelo,M., Collins,S., Camarata,J., Campopiano,A., Choepe,Y., Colangelo,M., Collins,S.,

Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fero,S., Ferreira,P., Fitzhugh,M., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Hearford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Karatas,A., Lacroque,K., Lamazares,R., Lander,T., Lenoczky,J., Levine,R., Liu,G., Maclean,C., MacDonald,P., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrum,J., Menes,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., Pollara,V., Raymond,C., Retta,R., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Ryback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Roselli,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Sougnuez,C., Spencer,B., Strange-Thomann,N., Stojanovic,N., Strassman,N., Subramanian,A., Talmas,J., Tesfaye,S., Theodore,J., Travers,M., Travis,N., Triggillo,J., Vassiliev,H., Videl,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
 TITLE Direct Submission
 JOURNAL Submitted (17-FEB-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 REFERENCE 3 (bases 1 to 166793)
 AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barnes,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fero,S., Ferreira,P., Fitzhugh,M., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamal,A., Karatas,A., Kells,C., Lacroque,K., Lamazares,R., Landers,T., Lenoczky,J., Levine,R., Liu,G., Maclean,C., MacDonald,P., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrum,J., Menes,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Ryback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Roselli,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Sougnuez,C., Spencer,B., Strange-Thomann,N., Stojanovic,N., Strassman,N., Subramanian,A., Talmas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Triggillo,J., Vassiliev,H., Videl,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
 TITLE Direct Submission
 JOURNAL Submitted (10-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 COMMENT On Apr 10, 2002 this sequence version replaced gi:19683210.
 All repeats were identified using RepeatMasker:
 Smit,A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/MIT Center for Genome Research
 Center code: WtBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 Project Information
 Center project name: L12709
 Center clone name: 729_B.4

NOTE: This is a 'working draft' sequence. It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
 This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.
 1 166793: contig of 166793 bp in length.
 Location/Qualifiers
 1..166793
 /organism="Homo sapiens"

THIS PAGE BLANK (USPTO)

XX WPI: 2001-408646/43.
 DR P-PSDB: AAU01210.
 XX Polynucleotide encoding novel human membrane protein, useful for
 PT identifying agonist, antagonist or modifiers or for producing
 PT antibodies useful in therapeutic, diagnostic and pharmacogenomic
 PT applications
 XX
 PS Claim 1: Page 29; 32pp; English.
 XX
 CC The present sequence encodes for novel human membrane protein #1.
 CC Human membrane protein #1 is 1 of 4 human membrane proteins
 CC (AAU01210-AAU01213) given in the present invention. These membrane
 CC proteins share structural similarity with membrane receptors such as
 CC the IGF receptor and mammalian CD20. The novel human membrane proteins
 CC are useful for identifying agonists, antagonists and modulators of the
 CC membrane proteins, and for producing antibodies specific to the
 CC membrane proteins. The membrane proteins can be used for diagnosis,
 CC drug screening, pharmacogenomic applications, clinical trial monitoring
 CC and the treatment of physiological disorders and diseases. The
 CC polynucleotides encoding the membrane proteins can be used to generate
 CC PCR primers or probes to identify mutations associated with a particular
 CC disease.
 CC
 SQ Sequence 603 BP; 162 A; 125 C; 104 G; 212 T; 0 other;

Query Match 100.0%; Score 603; DB 22; Length 603;
 Best Local Similarity 100.0%; Pred. No. 3,5e-152;
 Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGATTCAAGCAGCCGACAGAGTCGGTTCGTGATTTCTCCAGAAATCACTGCT 60
 DB 1 ATGGATTCAAGCAGCCGACAGAGTCGGTTCGTGATTTCTCCAGAAATCACTGCT 60
 QY 61 TCAGAAATAGAGTCCACAGAACTTTACGACGACCTTTTCAATCAAGCCCTTGCAA 120
 DB 61 TCAGAAATAGAGTCCACAGAACTTTACGACGACCTTTTCAATCAAGCCCTTGCAA 120
 QY 121 AAATTTTGTCTAGAAAATGAATCTTAGGACATACACAGATCCTGTTGGAAATTAAG 180
 DB 121 AAATTTTGTCTAGAAAATGAATCTTAGGACATACACAGATCCTGTTGGAAATTAAG 180
 QY 181 ACCTTTCTTTTGAGATATCTCTCTTTCACCTGTTTAAACCATATCCAGAGTTTCCC 240
 DB 181 ACCTTTCTTTTGAGATATCTCTCTTTCACCTGTTTAAACCATATCCAGAGTTTCCC 240
 Y 241 TTTATATTTCTTTCAGAGATATCCATCTGGGCTCTGTTTGTTCATTAATCTGAGCC 300
 DB 241 TTTATATTTCTTTCAGAGATATCCATCTGGGCTCTGTTTGTTCATTAATCTGAGCC 300
 QY 301 TTCCCTAATTCGAGTGAAGAAAACACAGAAACCTGTATATATTTAGCCGAATTAAG 360
 DB 301 TTCCCTAATTCGAGTGAAGAAAACACAGAAACCTGTATATATTTAGCCGAATTAAG 360
 QY 361 AATCTCTAGAGCCCTGAGAGCAATGCTGGAATCATCTCTCCATTTGGTTTCATC 420
 DB 361 AATCTCTAGAGCCCTGAGAGCAATGCTGGAATCATCTCTCCATTTGGTTTCATC 420
 QY 421 CTAGATCAAAACTACATTTGTGTTATTTCTACACCAAAATAGTCAGTGAAGGCTGTACT 480
 DB 421 CTAGATCAAAACTACATTTGTGTTATTTCTACACCAAAATAGTCAGTGAAGGCTGTACT 480
 QY 481 GTCCTGTTCTGGGAATTTTGATATGATGACTTTAGCATTTATTAATTAATTCATT 540
 DB 481 GTCCTGTTCTGGGAATTTTGATATGATGACTTTAGCATTTATTAATTAATTCATT 540
 QY 541 TCTCTGCTTTCATATTTGGGGTCCACTCAGAGAGATTGATTTGAACAATGTGT 600
 DB 541 TCTCTGCTTTCATATTTGGGGTCCACTCAGAGAGATTGATTTGAACAATGTGT 600
 QY 601 TGA 603
 DB 601 TGA 603

DB 601 TGA 603

RESULT 2

AAU04283
 ID AAU04283 standard; DNA; 676 BP.

AAU04283;

26-SEP-2001 (first entry)

DNA sequence encoding novel human membrane protein.

Human; membrane protein; membrane receptor; IGF receptor; CD20;

physiological disorder; ds.

Homo sapiens.

WO200146417-A2.

28-JUN-2001.

12-DEC-2000; 2000MO-US3742.

22-DEC-1999; 99US-0171567.

(LEXI-) LEXICON GENETICS INC.

Walke DM, Turner CA.

WPI: 2001-408646/43.

Polynucleotide encoding novel human membrane protein, useful for
 identifying agonist, antagonist or modifiers or for producing
 antibodies useful in therapeutic, diagnostic and pharmacogenomic
 applications

Disclosure: Page 31-32; 32pp; English.

The present sequence encoding for a novel human membrane protein
 includes the flanking 5'- and 3'- sequences. Four amino acid sequences
 for novel human membrane proteins (AAU01210-AAU01213) are given in the
 present invention. These membrane proteins share structural similarity
 with membrane receptors such as the IGF receptor and mammalian CD20.
 The novel human membrane proteins are useful for identifying agonist,
 antagonist and modulators of the membrane proteins, and for producing
 antibodies specific to the membrane proteins. The membrane proteins can
 be used for diagnosis, drug screening, pharmacogenomic applications,
 clinical trial monitoring and the treatment of physiological disorders
 and diseases. The polynucleotides encoding the membrane proteins can be
 used to generate PCR primers or probes to identify mutations associated
 with a particular disease.

Sequence 676 BP; 193 A; 140 C; 112 G; 231 T; 0 other;

Query Match 100.0%; Score 603; DB 22; Length 676;
 Best Local Similarity 100.0%; Pred. No. 3,7e-152;
 Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGATTCAAGCAGCCGACAGAGTCGGTTCGTGATTTCTCCAGAAATCACTGCT 60
 DB 33 ATGGATTCAAGCAGCCGACAGAGTCGGTTCGTGATTTCTCCAGAAATCACTGCT 92
 QY 61 TCAGAAATAGAGTCCACAGAACTTTACGACGACCTTTTCAATCAAGCCCTTGCAA 120
 DB 93 TCAGAAATAGAGTCCACAGAACTTTACGACGACCTTTTCAATCAAGCCCTTGCAA 152
 QY 121 AAATTTTGTCTAGAAAATGAATCTTAGGACATACACAGATCCTGTTGGAAATTAAG 180
 DB 153 AAATTTTGTCTAGAAAATGAATCTTAGGACATACACAGATCCTGTTGGAAATTAAG 212
 QY 181 ACCTTTCTTTTGAGATATCTCTCTTTCACCTGTTTAAACCATATCCAGAGTTTCCC 240
 DB 181 ACCTTTCTTTTGAGATATCTCTCTTTCACCTGTTTAAACCATATCCAGAGTTTCCC 240

```
Db 213 ACCTTTCTTTGGAGTTATCTTCCTTTACCTGTTAAACCATATCCAGGTTTCCC 272
Qy 241 TTTATATTTTTCAGAGATATCCATTTCTGGGGCTCTGTTTGTTCATTAATTCGAGCC 300
Db 273 TTTATATTTTTCAGAGATATCCATTTCTGGGGCTCTGTTTGTTCATTAATTCGAGCC 332
Qy 301 TTCCCTAATGACATGAAAAAAGAAAAACACAGAAACCTGATATATTTGACCAATTAAG 360
Db 333 TTCTTAATTCAGATGAAAAAAGAAAAACACAGAAACCTGATATATTTGAGCCGAATTAAG 392
Qy 361 AATCTTCTTAGTGCCTCGAGAGCAATAGCTGGAATCATCTCCTCAGATTTGGTTTCATC 420
Db 393 AATCTTCTTAGTGCCTCGAGAGCAATAGCTGGAATCATCTCCTCAGATTTGGTTTCATC 452
Qy 421 CTAGATCAAAATGACATTTGTGTTATTTCTACCCAAATATAGTCAGGTAAAGCTGTACT 480
Db 453 CTAGATCAAAATGACATTTGTGTTATTTCTACCCAAATATAGTCAGGTAAAGCTGTACT 512
Qy 481 GTCTGCTTTCTGGGAATTTTGAATTAATGATGATGATTTTCAGATTAATTAATTCATT 540
Db 513 GTCTGCTTTCTGGGAATTTTGAATTAATGATGATGATTTTCAGATTAATTAATTCATT 572
Qy 541 TCTCTGCTTTCTGCAATTTTGGGGTGCACACTCAGAGATTTGTGATTTGTAACAATGTGT 600
Db 573 TCTCTGCTTTCTGCAATTTTGGGGTGCACACTCAGAGATTTGTGATTTGTAACAATGTGT 632
Qy 601 TGA 603
Db 633 TGA 635

RESULT 3
AAH64741
ID AAH64741 standard; cDNA: 747 BP.
AC AAH64741;
XX
DT 11-SEP-2001 (first entry)
XX
DE Human secreted protein cDNA, SEQ ID NO: 17.
XX
KW Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;
KM GENSET; ss.
XX
OS Homo sapiens.
XX
PN WO200142451-A2.
XX
PD 14-JUN-2001.
XX
PF 07-DEC-2000: 2000MO-IB01938.
XX
PR 08-DEC-1999: 99US-0169629.
XX
PR 06-MAR-2000: 2000US-0187470.
XX
PA (GENSET) GENSET.
XX
PI Dumas Milne Edwards J, Bougueleret L, Jobert S;
XX
DR WPI: 2001-367870/38.
XX
DR P-PSDB: AAG89138.
XX
PT Full length GENSET human nucleic acids encoding potentially secreted
PT proteins, useful in gene therapy and vaccination against a variety of
PT diseases, and for diagnosis of those diseases -
XX
XX
XX Claim 7: Page 576-577; 921pp; English.
XX
XX
XX The invention relates to full length GENSET human nucleic acids encoding
XX potentially secreted proteins. The nucleic acids and the polypeptides
XX they encode may be used in the prevention, treatment and diagnosis of
XX diseases associated with inappropriate GENSET gene expression. For
XX example, they be used to treat disorders associated with decreased
```

```
CC GENSET gene expression by rectifying mutations or deletions in a
CC patient's genome that affect the activity of GENSET or by supplementing
CC the patient's own production of GENSET polypeptides. Conversely,
CC antisense nucleic acid molecules may be administered to down regulate
CC GENSET expression by binding with the cells' own genes and preventing
CC their expression. The sense and antisense nucleic acids may also be
CC used as DNA probes in diagnostic assays to detect and quantitate the
CC presence of similar nucleic acid sequences in samples, and hence to
CC determine which patients may be in need of restorative therapy.
CC The GENSET polypeptides may be used as antigens in the production of
CC antibodies and in assays to identify modulators (agonists and
CC antagonists) of GENSET polypeptide expression and activity. The
CC present sequence is a GENSET nucleic acid of the invention.
XX
SQ Sequence 747 BP; 240 A; 145 C; 121 G; 241 T; 0 other;
Qy Query Match 99.2%; Score 598.2; DB 22; Length 747;
Best Local Similarity 99.5%; Pred. No. 7.4e-151;
Matches 600; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 ATGATTCAGACACCGCACAGTCCGGTGTTCGTGATTTCTCCAGAAATCACTGCT 60
Db 73 ATGATTCAGACACCGCACAGTCCGGTGTTCGTGATTTCTCCAGAAATCACTGCT 132
Qy 61 TCAGAAATATGAGTCACAGAACTTTCAGCCAGACCTTTTCAACTCAAGCCCTTGCAA 120
Db 133 TCAGAAATATGAGTCACAGAACTTTCAGCCAGACCTTTTCAACTCAAGCCCTTGCAA 192
Qy 121 AATATATTTGCTAGAAAAATGAAAAATCTTAGGACATCCAGATCCCTGTTGGAATTAG 180
Db 193 AATATATTTGCTAGAAAAATGAAAAATCTTAGGACATCCAGATCCCTGTTGGAATTAG 252
Qy 181 ACCTTTCTTTTGGAGTATCTTCTTTTCACTTGTAAACCAATATCCAGAGTTCCC 240
Db 253 ACCTTTCTTTTGGAGTATCTTCTTTTCACTTGTAAACCAATATCCAGAGTTCCC 312
Qy 241 TTTATATTTCTTACAGATATCCATTCCTGGGGCTGTTTGTTCATTAATTCGAGCC 300
Db 313 TTTATATTTCTTACAGATATCCATTCCTGGGGCTGTTTGTTCATTAATTCGAGCC 372
Qy 301 TTCCCTAATGACATGAAAAAAGAAAAACACAGAAACCTGATTAATTTGAGCCGAATTAAG 360
Db 373 TTCCCTAATGACATGAAAAAAGAAAAACACAGAAACCTGATTAATTTGAGCCGAATTAAG 432
Qy 361 AATCTTCTTAGTGCCTCGAGAGCAATAGCTGGAATCATCTCCTCAGATTTGGTTTCATC 420
Db 433 AATCTTCTTAGTGCCTCGAGAGCAATAGCTGGAATCATCTCCTCAGATTTGGTTTCATC 492
Qy 421 CTAGATCAAAATGACATTTGTGTTATTTCTACCCAAATATAGTCAGGTAAAGCTGTACT 480
Db 493 CTAGATCAAAATGACATTTGTGTTATTTCTACCCAAATATAGTCAGGTAAAGCTGTACT 552
Qy 481 GTCTGCTTTCTGGGAATTTTGAATTAATGATGATGATTTTCAGATTAATTAATTCATT 540
Db 553 GTCTGCTTTCTGGGAATTTTGAATTAATGATGATGATTTTCAGATTAATTAATTCATT 612
Qy 541 TCTCTGCTTTCTGCAATTTTGGGGTGCACACTCAGAGATTTGTGATTTGTAACAATGTGT 600
Db 613 TCTCTGCTTTCTGCAATTTTGGGGTGCACACTCAGAGATTTGTGATTTGTAACAATGTGT 672
Qy 601 TGA 603
Db 673 TGA 675

RESULT 4
AAD21441
ID AAD21441 standard; cDNA: 760 BP.
AC AAD21441;
XX
XX
XX 28-JAN-2002 (first entry)
XX
```

DE Human CD20/IgE-receptor like protein encoding cDNA, agp-96614-al.
 XX
 XX Human; CD20/IgE-receptor like protein; immunoglobulin E; agp-96614-al;
 KW agp-69406-al; cancer; abnormal cell proliferation; autoimmune disease;
 KW ovarian cancer; brain cancer; arteriosclerosis; vascular stenosis;
 KW rheumatoid arthritis; multiple sclerosis; allergy; dermatitis; asthma;
 KW reproductive disease; diabetes; transplant rejection; endometriosis;
 KW infertility; gene therapy; ss.
 XX
 XX Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT CDS 98..700
 FT /tag= a
 FT /product= "Human CD20/IgE-receptor like protein,
 FT agp-96614-al."
 XX
 XX MO200174903-A2.
 XX
 XX 11-OCT-2001.
 XX
 XX 29-MAR-2001; 2001WO-US10048.
 XX
 XX 30-MAR-2000; 2000US-193728P.
 XX
 XX 27-NOV-2000; 2000US-0723258.
 XX
 XX (AMGE-) AMGEN INC.
 XX
 XX Welcher AA, Calzone FJ;
 XX WPI; 2001-662968/76.
 XX P-PSDB; AAE13062.
 XX
 XX Novel CD20/IgE-receptor like polypeptides and polynucleotides.
 PT antagonists and antibodies of the polypeptide useful for treating
 PT ameliorating or preventing diseases associated with the polypeptide
 PT e.g. cancer, asthma
 XX
 XX Claim 1; Fig 1; 145pp; English.
 PS
 XX
 XX The invention relates to human CD20/immunoglobulin E (IgE)-receptor
 CC like polypeptides designated as agp-96614-al and agp-69406-al and
 CC nucleic acid molecules encoding such polypeptides. Polypeptides of
 CC the invention are useful for treating, preventing or ameliorating
 CC a disease, condition, or disorder which includes cancer such as
 CC brain cancer, ovarian cancer; abnormal cell proliferation such as
 CC arteriosclerosis, vascular stenosis; pathology from allergens
 CC such as allergies, asthma, dermatitis; dysfunction of immune system
 CC such as rheumatoid arthritis, autoimmune disease, multiple sclerosis,
 CC diabetes, transplant rejection and reproductive diseases such as
 CC infertility, preterm labour and delivery, endometriosis etc. They
 CC are also useful for identifying antagonists and as immunogens, for
 CC raising antibodies which may also be used to prevent, treat or
 CC diagnose a number of diseases and disorders. Polynucleotides of the
 CC invention are used to map the location of CD20/IgE-receptor like
 CC gene and related genes on chromosomes and as hybridisation probes.
 CC They are also useful in gene therapy. The present sequence is
 CC human CD20/IgE-receptor like protein encoding cDNA, agp-96614-al.
 XX
 XX
 SQ Sequence 760 BP; 232 A; 157 C; 126 G; 245 T; 0 other;
 Query Match 99.2%; Score 598.2; DB 22; Length 760;
 Best Local Similarity 99.5%; Pred. No. 7.4e-151;
 Matches 600; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 121 AAATATATTTGCTAGAAAAATGAAAATCTTAGGACTATCCAGATCCTGTTGGAAATATG 180
 |||
 DB 218 AAATATATTTGCTAGAAAAATGAAAATCTTAGGACTATCCAGATCCTGTTGGAAATATG 277
 QY 181 ACCTTTCTTTTGGAGTATATCTTCTTCCCTTACCTGTGTTAAACCAATCCAGGTTTCC 240
 |||
 DB 278 ACCTTTCTTTTGGAGTATATCTTCTTCCCTTACCTGTGTTAAACCAATCCAGGTTTCC 337
 QY 241 TTTATATTTCTTTGAGATATCCATCTGGGCTGCTGTTTGTCAATTAATCTGGAGCC 300
 |||
 DB 338 TTTATATTTCTTTGAGATATCCATCTGGGCTGCTGTTTGTCAATTAATCTGGAGCC 397
 QY 301 TTCTTAATTTGACAGTAAAAAGAAAAACACAGAACTGTGATATAATGACGGAATATG 360
 |||
 DB 398 TTCTTAATTTGACAGTAAAAAGAAAAACACAGAACTGTGATATAATGACGGAATATG 457
 QY 361 AATCTTCTTAGTCCCTGAGACCAATAGCTGGAATCATCTCTCAATTTGGTTTCATC 420
 |||
 DB 458 AATCTTCTTAGTCCCTGAGACCAATAGCTGGAATCATCTCTCAATTTGGTTTCATC 517
 QY 421 CTAGATCAAACTACATTTGTTGTTATCTCAACCAAAATAGTCAGTAAAGCTGTACT 480
 |||
 DB 518 CTAGATCAAACTACATTTGTTGTTATCTCAACCAAAATAGTCAGTAAAGCTGTACT 577
 QY 481 GTCTGTCTTCTGGGAATTTGATTAATGATGACTTTCAGCATTAATTAATTAATCA 540
 |||
 DB 578 GTCTGTCTTCTGGGAATTTGATTAATGATGACTTTCAGCATTAATTAATTAATCA 637
 QY 541 TCTGTGCTTTTCAATTTTGGGCTGCTCAAGAGATTTGATTTGTAACAATGTTGT 600
 |||
 DB 638 TCTGTGCTTTTCAATTTTGGGCTGCTCAAGAGATTTGATTTGTAACAATGTTGT 697
 QY 601 TGA 603
 |||
 DB 698 TGA 700
 RESULT 5
 AAD18275 standard; cDNA; 689 BP.
 XX
 XX AAD18275;
 AC
 XX
 XX 18-DEC-2001 (first entry)
 DE
 XX Human immune system-related protein-encoding gene 9 cDNA clone HTENNA5.
 XX
 XX Human; immune system-related protein; allergy; rheumatoid arthritis;
 KW cancer; Gaucher's disease; viral hepatitis; gene therapy; cytostatic;
 KW diabetes mellitus; arhythmia; wound healing; ischaemic lesion; AIDS;
 KW Acquired Immune Deficiency Syndrome; virucide; hepatotropic; vasotropic;
 KW autoimmune disorder; inflammation; cardiovascular disorder; hair loss;
 KW wound healing; cell proliferation; skin aging; endocrine disorder;
 KW food preservative; ss.
 XX
 XX Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT CDS 29..634
 FT /tag= a
 FT /product= "Human immune-system related protein"
 XX
 XX MO20016722-A1.
 XX
 XX 13-SEP-2001.
 XX
 XX 07-MAR-2001; 2001WO-US07260.
 XX
 XX 08-MAR-2000; 2000US-187873P.
 XX
 XX 11-AUG-2000; 2000US-224367P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX

PI NI J, Hilbert D, Kenny JJ, Moore PA, Choi GH, Soppet DR, Ebner R;
PI Gruber JR, Endress GA, Ruben SM;
XX WPI: 2001-569939/66.
DR P-PSDB: AAE10917.
XX
XX Novel isolated immune system-related polypeptide useful for treating
PT rheumatoid arthritis, AIDS, allergy, cancer, Gaucher's disease,
PT diabetes mellitus, arrhythmia, wound healing, ischemic lesions and
PT viral hepatitis -
PS
XX
PS Claim 1: Page 301; 315bp; English.

CC The invention relates to human immune system-related protein and their
CC DNA. Human immune-system related protein and DNA are useful for
CC preventing, treating or ameliorating a medical condition in a mammalian
CC subject, for diagnosing, preventing or treating immune system-associated
CC disorders, autoimmune disorders (rheumatoid arthritis), inflammatory
CC disorders (allergies), immunodeficiencies (AIDS), inflammatory disorders
CC (allergies), infectious diseases (e.g., viral hepatitis), complement
CC activation disorders, immune complex diseases, neoplastic disorders
CC (cancer), hyperproliferative disorders (Gaucher's disease), disorders
CC associated with neovascularisation, diseases at the cellular level,
CC cardiovascular disorders (arrhythmias), wound healing and epithelial
CC cell proliferation, endocrine disorders (diabetes mellitus) and
CC neurological disorders (ischemic lesions). Immune-system related protein
CC or DNA is useful for preventing hair loss, skin aging due to sunburn, to
CC maintain organs before transplantation, to treat weight disorders, to
CC modulate mammalian characteristics, to change a mammal's mental or
CC physical state, or as a food additive or preservative. Immune-system
CC related DNA is useful in gene therapy, for chromosome identification,
CC radiation hybrid mapping, long range restriction mapping and in forensic
CC biology. The present sequence represents a human immune-system related
CC protein-encoding cDNA of the invention.

XX
XX Sequence 689 BP; 204 A; 138 C; 114 G; 233 T; 0 other;

Query Match 97.5%; Score 587.8; DB 22; Length 689;
Best Local Similarity 99.5%; Pred. No. 4.5e-148;

Matches 600; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 ATGGATTCAAGCACCGACAGTCGGGTGTTCTGGTATTTCCTCAGAAATCACTGCT 60
DB 29 ATGGATTCAAGCACCGACAGTCGGGTGTTCTGGTATTTCCTCAGAAATCACTGCT 88
QY 61 TCAGAAATAGATCAGACAGAACTTTCAGCCAGACCTTTCAACTCAAGCCCTTGCAA 120
DB 89 TCAGAAATAGATCAGACAGAACTTTCAGCCAGACCTTTCAACTCAAGCCCTTGCAA 148
QY 121 AAATATTATTCGTGAAGAAAGAAATCTAGGGACTATCCAGATCCCTGTTGGAATATG 180
DB 149 AAATATTATTCGTGAAGAAAGAAATCTAGGGACTATCCAGATCCCTGTTGGAATATG 208
QY 181 ACCTTTCTTTGGAGATTATCTCTTTTCACCTGTTTAAACCATATCCAAAGGTTTCCC 240
DB 209 ACCTTTCTTTGGAGATTATCTCTTTTCACCTGTTTAAACCATATCCAAAGGTTTCCC 268
QY 241 TTTATATTCTTTCAGAGATATCCATCTGGGGCTCTGTTTGTTCATTTAATTCGAGCC 300
DB 269 TTTATATTCTTTCAGAGATATCCATCTGGGGCTCTGTTTGTTCATTTAATTCGAGCC 328
QY 301 TTCTATATTCGAGTGAAGAAAGAAAGAAACACAGAAATCTGATTAATATGAGCCGAATATG 360
DB 329 TTCTATATTCGAGTGAAGAAAGAAAGAAACACAGAAATCTGATTAATATGAGCCGAATATG 388
QY 361 AATCTCTTAGTGCCCTGAGAGCAATAGTGAATCAATTCCTCCACATTTGGTTTCATC 420
DB 389 AATTTTCTTAGTGCCCTGAGAGCAATAGTGAATCAATTCCTCCACATTTGGTTTCATC 448
QY 421 CTAGATCAAAACTACATTTTGGTTATTCACCAAAATAGTCAAGTGAAGCTGTACT 480
DB 449 CTAGATCAAAACTACATTTTGGTTATTCACCAAAATAGTCAAGTGAAGCTGTACT 508

QY 481 GTCTGTTCCTTGGGAATTTTGATTACATGTAGTACTTTCAGCAATTAATTAATTCATT 540
DB 509 GTCTGTTCCTTGGGAATTTTGATTACATGTAGTACTTTCAGCAATTAATTAATTCATT 568
QY 541 TCTCTGCCCTTTCATTAATTTGGGGTGCACACTCAGAGGATTTGTGTAACAATGTTGT 600
DB 569 TCTCTGCCCTTTCATTAATTTGGGGTGCACACTCAGAGGATTTGTGTAACAATGTTGT 627
QY 601 TGA 603
DB 628 TGA 630

RESULT 6

AL01084/c
ID AL01084 standard; cDNA: 737 BP.
XX
XX
XX AL01084;
XX
XX 21-NOV-2001 (first entry)
XX
XX Human reproductive system related antigen cDNA SEQ ID NO: 1085.
DE
XX Human: reproductive system related antigen; reproductive system disorder;
KW cancer; gene therapy; ss.
OS
XX Homo sapiens.
OS
XX WO200155320-A2.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01339.
XX
XX 31-JAN-2000; 2000US-0179065.
XX 04-FEB-2000; 2000US-0180628.
XX 24-FEB-2000; 2000US-0184664.
XX 02-MAR-2000; 2000US-0186350.
XX 16-MAR-2000; 2000US-0189874.
XX 17-MAR-2000; 2000US-0190076.
XX 18-APR-2000; 2000US-0198123.
XX 19-MAY-2000; 2000US-0205515.
XX 07-JUN-2000; 2000US-0209467.
XX 28-JUN-2000; 2000US-0214886.
XX 30-JUN-2000; 2000US-0215135.
XX 07-JUL-2000; 2000US-0216647.
XX 07-JUL-2000; 2000US-0216880.
XX 11-JUL-2000; 2000US-0217487.
XX 11-JUL-2000; 2000US-0217496.
XX 14-JUL-2000; 2000US-0218390.
XX 26-JUL-2000; 2000US-0220963.
XX 26-JUL-2000; 2000US-0220964.
XX 14-AUG-2000; 2000US-0224518.
XX 14-AUG-2000; 2000US-0224519.
XX 14-AUG-2000; 2000US-0225213.
XX 14-AUG-2000; 2000US-0225214.
XX 14-AUG-2000; 2000US-0225266.
XX 14-AUG-2000; 2000US-0225267.
XX 14-AUG-2000; 2000US-0225268.
XX 14-AUG-2000; 2000US-0225270.
XX 14-AUG-2000; 2000US-0225477.
XX 14-AUG-2000; 2000US-0225478.
XX 14-AUG-2000; 2000US-0225757.
XX 14-AUG-2000; 2000US-0225758.
XX 14-AUG-2000; 2000US-0225759.
XX 18-AUG-2000; 2000US-0226279.
XX 22-AUG-2000; 2000US-0226681.
XX 22-AUG-2000; 2000US-0226682.
XX 22-AUG-2000; 2000US-0227182.
XX 23-AUG-2000; 2000US-0227009.
XX 30-AUG-2000; 2000US-0228924.
XX 01-SEP-2000; 2000US-0229287.
XX 01-SEP-2000; 2000US-0229343.
XX 01-SEP-2000; 2000US-0229344.

OY 361 AATCTCTAGTGGCCCTGAGACAAATAGCTGGAATTCCTCTACATTTGGTTTCATC 420
|||
Db 326 AATTTCTTATGCCCTGGAGCAATAGCTGGAATTCATCTCTCCACATTTGGTTTCATC 267
OY 421 CTAGATCAAAACATCATTTTGGTTATTTCTACCAAAATAGTCAAGTGAAGCGTTACT 480
|||
Db 266 CTAGATCAAAACATCATTTTGGTTATTTCTACCAAAATAGTCAAGTGAAGCGTTACT 207
OY 481 GTCCTGTTCTTGGCAATTTTGATTCATTTGATGACTTTCGCAATTTATTTATTCATT 540
|||
Db 206 GTCCTGTTCTTGGCAATTTTGATTCATTTGATGACTTTCGCAATTTATTTATTCATT 147
OY 541 TCTCTGCTTCTTCATATTTTGGGTCCTCAGAGGATTTGTGATTTGAACATTTGT 600
|||
Db 146 TCTCTGCTTCTTCATATTTTGGGTCCTCAGAGGATTTGTGATTTGAACATTTGT 88
OY 601 TGA 603
|||
Db 87 TGA 85

RESULT 7
ABL96550/c
ID ABL96550 standard; cDNA; 737 BP.
XX
AC ABL96550;
XX
DT 21-JUN-2002 (first entry)
XX
DE Human testicular antigen encoding cDNA SEQ ID NO: 218.
XX
KW Human; testicular antigen; testes; cancer; metastasis; immune disorder;
KW reproductive system disorder; urinary system disorder; gene therapy;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disease; infection; cytostatic; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200155317-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01329.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.

PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234423.
PR 21-SEP-2000; 2000US-0234424.
PR 25-SEP-2000; 2000US-0234597.
PR 25-SEP-2000; 2000US-0234598.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.

XX Human proteins and DNA encoding sequences useful for preventing,
 PT treating or ameliorating a medical condition in a mammalian subject
 PT e.g. arthritis and cancer -

XX Claim 1: Page 961; 1963pp; English.

XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
 CC sequences ABA08225-ABA0574 represent nucleic acids encoding them. The
 CC invention also relates to vectors and recombinant host cells comprising a
 CC nucleotide of the invention, methods of producing the novel polypeptides,
 CC antibodies against the polypeptides, methods of detecting the nucleotides
 CC or polypeptides in a sample, and methods of identifying compounds which
 CC bind to polypeptides of the invention. Although novel, many of the
 CC polypeptides of the invention have homology to known proteins, thereby
 CC giving an insight into their probable biological activities, and hence
 CC potential therapeutic applications. The polypeptides of the invention may
 CC have various activities, including cytokine, cell proliferation or cell
 CC differentiation activities; stem cell growth factor activity;
 CC haematopoiesis regulatory activity; tissue growth factor activity;
 CC immunomodulatory activity; activin- or inhibin-related activities;
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
 CC thrombolytic activities; receptor or ligand activities; or may be
 CC involved in oncogenesis, cancer cell proliferation or metastasis.
 CC Depending on their biological activities, polypeptides and nucleotides of
 CC the invention are useful for preventing, treating or ameliorating medical
 CC conditions, e.g., by protein or gene therapy. Such conditions include
 CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
 CC vascular growth. Polypeptides involved with tissue regeneration and
 CC repair (or nucleic acids encoding them) may be used to promote wound
 CC healing (e.g., of burns, incisions and ulcers), while those with
 CC immunomodulatory activities may be used in the treatment of viral,
 CC bacterial and fungal infections in addition to immune disorders.
 CC Polypeptides with growth factor activity may be used in cell cultures to
 CC promote cell growth. For example, such polypeptides may be used to
 CC manipulate stem cells in culture to give rise to neuroepithelial cells
 CC that can be used to augment or replace cells damaged by illness,
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides
 CC may also be used in the diagnosis of the above conditions, and in drug
 CC screening techniques. The present sequence represents a cDNA encoding a
 CC novel human polypeptide of the invention.

XX Sequence 695 BP; 203 A; 142 C; 116 G; 234 T; 0 other;

XX Query Match 90.8%; Score 547.8; DB 22; Length 695;

XX Best Local Similarity 98.2%; Pred. No. 2.5e-137;

XX Matches 596; Conservative 0; Mismatches 7; Indels 4; Gaps 4;

QY 1 ATGATTTCAAGACCGCACACAGTCCGGTGTTCGTATTCCTCCAGAAATCACTGCT 60
 DB 52 ATGATTTCAAGACCGCACACAGTCCGGTGTTCGTATTCCTCCAGAAATCACTGCT 111
 QY 61 TCAGAAATTCAGTCCACAGAACTTCAGCAGACCTTTCAACCAAGCCCTTGCAA 120
 DB 112 TCAGAAATTCAGTCCACAGAACTTCAGCAGACCTTTCAACCAAGCCCTTGCAA 171
 QY 121 AAATTTATTTGCTAGAAAATGAAATCTTAGGACTATCCAGATCCTGTTGGAATTATG 180
 DB 172 AAATTTATTTGCTAGAAAATGAAATCTTAGGACTATCCAGATCCTGTTGGAATTATG 231
 QY 181 ACCTTTCTTTGGAGTATCTCTCTTTACCTGTTTAAACCAATCCAAAGGTTTCCC 240
 DB 232 ACCTTTCTTTGGAGTATCTCTCTTTACCTGTTTAAACCAATCCAAAGGTTTCCC 291
 QY 241 TTATATTTCTTCAGATATCATTCGCGGCTGTTTGTTCATTAATTCGAGACC 300
 DB 292 TTATATTTCTTCAGATATCATTCGCGGCTGTTTGTTCATTAATTCGAGACC 351
 QY 301 TTCCTAATTCGAGTAAAAAGAAAACACAGAACTGATTAATTTAGCCGATTAATG 360
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

DB 352 TTCCTAATTCGAGTAAAAAGAAAACACAGAAACTCGATTAATTTAGCCGAAATATG 411
 QY 361 AATCTTCTTAGAGCCCTGAGACCAATAGCTGGAATCATCTCTCATATTG-CTTTCAT 419
 DB 412 AATTTCTTAGAGCCCTGAGACCAATAGCTGGAATCATCTCTCATATTGATTTTCAT 471
 QY 420 CCTAGTCAAACTCATTTGTTGTTATTCACCAAAATATGATGTTAAGGCTGTAC 479
 DB 472 CCTAGTCAAACTCATTTGTTGTTATTCACCAAAATATGATGTTAAGGCTGTAC 531
 QY 480 TGTCTGTTCTTGGGAATTTGATTAATTCATTCATTCATTCATTCATTCATTCATTCAT 538
 DB 532 TGTCTGTTCTTGGGAATTTGATTAATTCATTCATTCATTCATTCATTCATTCATTCAT 591
 QY 539 T-TTCTTCGCTTT-CTCATTTTGGGGTCCACTCAGAGGATTTGATTTGAAACATG 596
 DB 592 TTTCTTCGCTTTACTCAATTTGGGGTCCACTCAGAGGATTTGATTTGAAACATG 651
 QY 597 TTGTTGA 603
 DB 652 TTGTTGA 658
 |||||||

RESULT 9

AA504282
 ID AA504282 standard; cDNA; 450 BP.

AA504282;

26-SEP-2001 (first entry)

Novel human membrane protein #4 cDNA sequence.

Human; membrane protein; membrane receptor; Ige receptor; CD20;
 physiological disorder; ss.

Homo sapiens.

Key Location/Qualifiers

CDS 1..450
 /*tag= a
 /product= "Human membrane protein #4"

W0200146417-A2.

28-JUN-2001.

12-DEC-2000; 2000WO-US33742.

22-DEC-1999; 99US-0171567.

(LEXI-) LEXICON GENETICS INC.

Walke DW, Turner CA;

WPI: 2001-408646/43.

P-PSDB; AAU01213.

Polynucleotide encoding novel human membrane protein, useful for
 identifying agonist, antagonist or modifiers or for producing
 antibodies useful in therapeutic, diagnostic and pharmacogenomic
 applications -

Disclosure: Page 31; 32pp; English.

The present sequence encodes for novel human membrane protein #4.

Human membrane protein #4 is 1 of 4 human membrane proteins
 (AAU01210-AAU01213) given in the present invention. These membrane
 proteins share structural similarity with membrane receptors such as
 the Ige receptor and mammalian CD20. The novel human membrane proteins
 are useful for identifying agonists, antagonists and modulators of the
 membrane proteins, and for producing antibodies specific to the
 membrane proteins. The membrane proteins can be used for diagnosis.

CC drug screening, pharmacogenomic applications, clinical trial monitoring
CC and the treatment of physiological disorders and diseases. The
CC polynucleotides encoding the membrane proteins can be used to generate
CC PCR primers or probes to identify mutations associated with a particular
CC disease.

XX Sequence 450 BP; 120 A; 94 C; 77 G; 159 T; 0 other;

Query Match 57.6%; Score 347.4; DB 22; Length 450;
Best Local Similarity 97.0%; Pred. No. 1.4e-83;
Matches 354; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

OY 1 ATGATTCAGACGCGACACAGTCGGTGTTCGTATTTCTCCAGAAATCACTGCT 60
DB 1 ATGATTCAGACGCGACACAGTCGGTGTTCGTATTTCTCCAGAAATCACTGCT 60
OY 61 TCAGAAATATGAGTCCACAGAACTTTCAGCCAGACCTTTCAACGCCCCCTTGCAA 120
DB 61 TCAGAAATATGAGTCCACAGAACTTTCAGCCAGACCTTTCAACGCCCCCTTGCAA 120
OY 121 AAATATTTTGTAGAAAATGAAAATCTTAGGACTATCCAGATCTGTTGGAATTATG 180
DB 121 AAATATTTTGTAGAAAATGAAAATCTTAGGACTATCCAGATCTGTTGGAATTATG 180
OY 181 ACCTTTCTTTTGGAGTATCTTCCTTTTCACCTTTTAAACATATCCAAAGTTTCCC 240
DB 181 ACCTTTCTTTTGGAGTATCTTCCTTTTCACCTTTTAAACATATCCAAAGTTTCCC 240
OY 241 TTTATATTTCTTTCAGGATATCCATTCGGGCTCTGTTTGTATTAATCTTGAGCC 300
DB 241 TTTATATTTCTTTCAGGATATCCATTCGGGCTCTGTTTGTATTAATCTTGAGCC 300
OY 301 TTCCATATTTGAGTGAAGAAAAGAAAACACAGAACTCTGATATATGACCGAATATG 360
DB 301 TTCCATATTTGAGTGAAGAAAAGAAAACACAGAACTCTGATATATGACCGAATATG 360
OY 361 AATCT 365
DB 361 ACTTT 365

RESULT 10
AAS04281
ID AAS04281 standard; cDNA; 417 BP.

XX AAS04281;

XX 26-SEP-2001 (first entry)

DE Novel human membrane protein #3 cDNA sequence.

XX Human; membrane protein; membrane receptor; Ige receptor; CD20;
KW physiological disorder; ss.

XX Homo sapiens.

XX Key Location/Qualifiers
FH 1..417
FT CDS /tag= a
FT /product= "Human membrane protein #3"

XX WO200146417-A2.

XX 28-JUN-2001.

XX 12-DEC-2000; 2000WO-US33742.

XX 22-DEC-1999; 99US-0171567.

XX (LEXI-) LEXICON GENETICS INC.

XX Walke DW, Turner CA;

DR WPI; 2001-408646/43.
DR P-PSDB; AAU01212.

XX Polynucleotide encoding novel human membrane protein, useful for
PT identifying agonist, antagonist or modulators or for producing
PT antibodies useful in therapeutic, diagnostic and pharmacogenomic
PT applications

XX Disclosure; Page 30; 32pp; English.

CC The present sequence encodes for novel human membrane protein #3.
CC Human membrane protein #3 is 1 of 4 human membrane proteins
CC (AAU01210-AAU01213) given in the present invention. These membrane
CC proteins share structural similarity with membrane receptors such as
CC the Ige receptor and mammalian CD20. The novel human membrane proteins
CC are useful for identifying agonists, antagonists and modulators of the
CC membrane proteins, and for producing antibodies specific to the
CC membrane proteins. The membrane proteins can be used for diagnosis,
CC drug screening, pharmacogenomic applications, clinical trial monitoring
CC and the treatment of physiological disorders and diseases. The
CC polynucleotides encoding the membrane proteins can be used to generate
CC PCR primers or probes to identify mutations associated with a particular
CC disease.

XX Sequence 417 BP; 115 A; 91 C; 69 G; 142 T; 0 other;

Query Match 56.7%; Score 341.8; DB 22; Length 417;
Best Local Similarity 99.4%; Pred. No. 4.5e-82;
Matches 343; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ATGATTCAGACGCGACACAGTCGGTGTTCGTATTTCTCCAGAAATCACTGCT 60
DB 1 ATGATTCAGACGCGACACAGTCGGTGTTCGTATTTCTCCAGAAATCACTGCT 60
OY 61 TCAGAAATATGAGTCCACAGAACTTTCAGCCAGACCTTTCAACGCCCCCTTGCAA 120
DB 61 TCAGAAATATGAGTCCACAGAACTTTCAGCCAGACCTTTCAACGCCCCCTTGCAA 120
OY 121 AAATATTTTGTAGAAAATGAAAATCTTAGGACTATCCAGATCTGTTGGAATTATG 180
DB 121 AAATATTTTGTAGAAAATGAAAATCTTAGGACTATCCAGATCTGTTGGAATTATG 180
OY 181 ACCTTTCTTTTGGAGTATCTTCCTTTTCACCTTTTAAACATATCCAAAGTTTCCC 240
DB 181 ACCTTTCTTTTGGAGTATCTTCCTTTTCACCTTTTAAACATATCCAAAGTTTCCC 240
OY 241 TTTATATTTCTTTCAGGATATCCATTCGGGCTCTGTTTGTATTAATCTTGAGCC 300
DB 241 TTTATATTTCTTTCAGGATATCCATTCGGGCTCTGTTTGTATTAATCTTGAGCC 300
OY 301 TTCCATATTTGAGTGAAGAAAAGAAAACACAGAACTCTGATATATA 345
DB 301 TTCCATATTTGAGTGAAGAAAAGAAAACACAGAACTCTGATATATA 345

RESULT 11

ABL67173/C
ID ABL67173 standard; DNA; 382 BP.

XX ABL67173;

XX 15-MAY-2002 (first entry)

XX Thyroid cancer related gene sequence SEQ ID NO:5510.

XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;
KW gene; ds.

XX Homo sapiens.

XX WO200194629-A2.

XX 13-DEC-2001.
PD
XX
PF 30-MAY-2001; 2001MO-US10838.
XX
XX 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237112P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
XX WPI; 2002-188264/24.
DR
XX
PT Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set -
XX
XX Claim 1; SEQ ID 5510; 44pp; English.
PS
XX
XX The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the

CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilms' tumour.
XX
SQ Sequence 382 BP; 135 A; 73 C; 67 G; 107 T; 0 other;
Query Match 56.4%; Score 340.2; DB 24; Length 382;
Best Local Similarity 99.1%; Pred. No. 1.2e-81;
Matches 342; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 259 FATCCATTCCTGGGCGCTCTGTTTGTTCATTAATCTGAGCCCTCCATATGCGAGTGA 318
|||||
DB 382 TATCCATTCCTGGGCGCTCTGTTTGTTCATTAATCTGAGCCCTCCATATGCGAGTGA 323
|||||
QY 319 AGAAAAACACAGAAACTCTGATTAATATGAGCGAATATGATCTTCTAGTGCCTG 378
|||||
DB 322 AGAAAAACACAGAAACTCTGATTAATATGAGCGAATATGATCTTCTAGTGCCTG 263
|||||
QY 379 AGAGCAATAGCTGGATCATCTTCCTCACATTTGGTTTCATCTAGATCAAACTACATT 438
|||||
DB 262 GGAGCAATAGCTGGATCATCTTCCTCACATTTGGTTTCATCTAGATCAAACTACATT 203
|||||
QY 439 TGTGGTTATCTCACCAAAATAGTCAGTGAAGGCTGTCTGCTGCTTCTGGGGAATT 498
|||||
DB 202 TGTGGTTATCTCACCAAAATAGTCAGTGAAGGCTGTCTGCTGCTTCTGGGGAATT 143
|||||
QY 499 TTGATTACATTTGATGACTTTTCACATTAATTAATTAATTTCTGCTTCTGCAATT 558
|||||
DB 142 TTGATTACATTTGATGACTTTTCACATTAATTAATTAATTTCTGCTTCTGCAATT 83
|||||
QY 559 TTGGGGTGGCCATCAGAGAGATTGTGATTTGGAACAATGTTGTGA 603
|||||
DB 82 TTGGGGTGGCCATCAGAGAGATTGTGATTTGGAACAATGTTGTGA 38
|||||
RESULT 12
AAC03080
ID AAC03080 standard; cDNA; 372 BP.
XX
AC AAC03080;
XX
XX 06-OCT-2000 (first entry)
DT
XX
DE Human secreted protein 5' EST, SEQ ID NO: 3078.
DE
XX
KW Human: 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
XX
PN BP1033401-A2.
XX
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 99US-0122487.
XX
XX (GENSET) GENSET.
PA
XX
XX Dumas Milne Edwards J, Duclert A, Giordano J;
PI
XX
XX WPI; 2000-500381/45.
DR
XX P-PSDB: AAG03074.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
XX Claim 1; SEQ ID 3078; 71pp + CD-ROM; English.

PR 08-DEC-1999; 99US-0169629.
 PR 06-MAR-2000; 2000US-0187470.
 XX (BEST) GENSET.
 PA
 XX Dumas Milne Edwards J, Bougueleret L, Jobert S;
 DR WPI: 2001-367870/38.
 DR P-PSDB; AAG89142.
 XX
 PT Full length GENSET human nucleic acids encoding potentially secreted
 PT proteins, useful in gene therapy and vaccination against a variety of
 PT diseases, and for diagnosis of those diseases -
 XX
 PS Claim 7; Page 581; 921pp; English.
 XX
 CC The invention relates to full length GENSET human nucleic acids encoding
 CC potentially secreted proteins. The nucleic acids and the polypeptides
 CC they encode may be used in the prevention, treatment and diagnosis of
 CC diseases associated with inappropriate GENSET gene expression. For
 CC example, they be used to treat disorders associated with decreased
 CC GENSET gene expression by rectifying mutations or deletions in a
 CC patient's genome that affect the activity of GENSET or by supplementing
 CC the patient's own production of GENSET polypeptides. Conversely,
 CC antisense nucleic acid molecules may be administered to down regulate
 CC GENSET expression by binding with the cells' own genes and preventing
 CC their expression. The sense and antisense nucleic acids may also be
 CC used as DNA probes in diagnostic assays to detect and quantitate the
 CC presence of similar nucleic acid sequences in samples, and hence to
 CC determine which patients may be in need of restorative therapy.
 CC The GENSET polypeptides may be used as antigens in the production of
 CC antibodies and in assays to identify modulators (agonists and
 CC antagonists) of GENSET polypeptide expression and activity. The
 CC present sequence is a GENSET nucleic acid of the invention.
 XX
 XX Sequence 468 BP; 155 A; 94 C; 81 G; 138 T; 0 other;
 S0
 Query Match 26.3%; Score 158.4; DB 22; Length 468;
 Best Local Similarity 75.0%; Pred. NO. 8.4e-33;
 Matches 199; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
 QY 1 ATGGATTCAAGCACACCGACAGTCCGGTGTCTGTGTAATTCCTCCAGAAATCACTGCT 60
 DB 76 ATGGATTCAAGCACACCGACAGTCCGGTGTCTGTGTAATTCCTCCAGAAATCACTGCT 135
 QY 61 TCAGAAATATGAGTCCACAGAACTTTCAGCCACGACTTTTCAACTCAAGCCCTTGCAA 120
 DB 136 TCAGAAATATGAGTCCACAGAACTTTCAGCCACGACTTTTCAACTCAAGCCCTTGCAA 195
 QY 121 AAATTAATTTCTGTAAGAAAATGTAAGGAAATCTTGGGACTATCCAGATCCCTTTGGAATATG 180
 DB 196 AAATTAATTTCTGTAAGAAAATGTAAGGAAATCTTGGGACTATCCAGATCCCTTTGGAATATG 255
 QY 181 ACCTTTCTTTTGGAGTTATCTCTTTTCACTTTTAAACCATATCCAGGTTTCCC 240
 DB 256 TCATTAATTTCTGGAGCTTCTCTTAATTCAGTAAGAAAACACACAGAAACTCTGGGAA 315
 QY 241 TTTATATTTTCTTTCAGGATATCCA 264
 DB 316 TTTTGATTACATGATGATCTTCA 339
 RESULT 15
 AAA16693
 ID AAA16693 standard; cDNA; 1330 BP.
 XX
 AC AAA16693;
 XX
 DT 16-JUN-2000 (first entry)
 XX
 XX Human secreted protein clone pe246_1 nucleotide sequence SEQ ID NO:151.
 DE Human; secreted protein; immunestimulant; immunosuppressant; virucide;
 XX
 KM

KM antibacterial; antifungal; cytostatic; antiinflammatory; dermatological;
 KM antidiabetic; antisthmatic; antiarthritic; antirheumatic; protozoacide;
 KM antihyroid; immune deficiency; severe combined immunodeficiency; SCID;
 KM infection; HIV; hepatitis; malaria; autoimmune disorder; systemic lupus;
 KM connective tissue disease; multiple sclerosis; erythematosis;
 KM rheumatoid arthritis; autoimmune pulmonary inflammation; asthma;
 KM Guillain-Barre syndrome; autoimmune thyroiditis; myasthenia gravis;
 KM insulin dependent diabetes mellitus; graft-versus-host disease;
 KM autoimmune inflammatory eye disease; allergy; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200009552-A1.
 XX
 PD 24-FEB-2000.
 XX
 PF 13-AUG-1999; 99WO-US18298.
 XX
 PR 14-AUG-1998; 98US-0096622.
 PR 17-AUG-1998; 98US-0096815.
 PR 04-SEP-1998; 98US-0099229.
 PR 23-OCT-1998; 98US-0105368.
 PR 08-JAN-1999; 99US-0115234.
 PR 12-FEB-1999; 99US-0119931.
 PR 18-FEB-1999; 99US-0120575.
 PR 30-APR-1999; 99US-0132020.
 PR 11-AUG-1999; 99US-0096622.
 XX
 PA (GEMY) GENETICS INST INC.
 XX
 PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
 PI Merberg D, Treacy M, Agostino MJ, Steininger RJ, Spaulding V;
 PI Wong GG, Clark HF, Fechtel K;
 XX
 DR WPI: 2000-205979/18.
 DR P-PSDB; AAY94973.
 XX
 PT New polynucleotides encoding secreted proteins, which may have e.g.
 PT nutritional, chemokine, immune stimulating or suppressing,
 PT hematopoiesis regulating, tissue growth, activin/inhibin
 PT antiinflammatory or tumor inhibition activity -
 XX
 PS Claim 160; Page 606-607; 641pp; English.
 XX
 CC AAA16618 to AAA16697 encode the human secreted proteins given in
 CC AAY94898 to AAY94980, isolated from human adult brain, adult thyroid,
 CC adult retina, foetal carcinoma, adult blood, adult neural, foetal kidney,
 CC adult placenta, adult testis, whole embryo, adult cartilage, kidney,
 CC foetal brain, adult thymus, foetal placenta, adult uterus' adult tumour,
 CC and adult bladder, cDNA libraries. The polynucleotides and proteins are
 CC predicted to have biological activities which would make them suitable
 CC for treating, preventing or ameliorating medical conditions in humans
 CC and animals. The polynucleotides can be used as markers for tissues in
 CC which the protein is preferentially expressed, as molecular weight
 CC markers on Southern gels, and as chromosome markers or tags to identify
 CC chromosomes or to map gene positions. The proteins can be used in the
 CC treatment of immune deficiencies and disorders, such as severe combined
 CC immunodeficiency (SCID), as well as viral, bacterial, fungal and other
 CC infections. These infections include human immunodeficiency virus (HIV),
 CC hepatitis, herpesviruses, mycobacteria, Leishmania spp., malaria and
 CC candidiasis. The proteins can be used to treat autoimmune disorders such
 CC as connective tissue disease, multiple sclerosis, systemic lupus
 CC erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation,
 CC Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent
 CC diabetes mellitus, myasthenia gravis, graft-versus-host-disease and
 CC autoimmune inflammatory eye disease. The proteins can also be used to
 CC treat allergic conditions, such as asthma. AAA16698 to AAA16774 represent
 CC probes for the human secreted proteins from the present invention.
 CC
 XX
 S0 Sequence 1330 BP; 364 A; 285 C; 309 G; 372 T; 0 other;
 Query Match 14.3%; Score 86.4; DB 21; Length 1330;
 Best Local Similarity 51.1%; Pred. NO. 2.6e-13;
 XX

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 22, 2003, 08:29:55 : Search time 54 Seconds

(without alignments)
3424.557 Million cell updates/sec

Title: US-09-735-712-1

Perfect score: 603
Sequence: 1 atgattcaacgacgcgcaca.....attgtgacaattgttga 603

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 segs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5a.COMB.seq: *
2: /cgn2_6/prodata/1/ina/5a.COMB.seq: *
3: /cgn2_6/prodata/1/ina/5a.COMB.seq: *
4: /cgn2_6/prodata/1/ina/5a.COMB.seq: *
5: /cgn2_6/prodata/1/ina/5a.COMB.seq: *
6: /cgn2_6/prodata/1/ina/5a.COMB.seq: *

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	86.4	14.3	1669	2	US-08-916-902A-2
2	86.4	14.3	1669	2	US-09-213-389-2
3	63.2	10.5	1661	1	US-08-318-492-3
4	63.2	10.5	1661	1	US-08-707-340-3
5	63.2	10.5	1661	1	US-08-994-578-3
6	52.6	8.7	945	2	US-09-149-476-168
7	49.2	8.2	2545	1	US-07-869-933-22
8	49.2	8.2	2545	1	US-09-103-663-22
9	48	8.0	1308	4	US-09-724-864-15
10	42.6	7.1	7218	1	US-08-232-463-14
11	40.8	6.8	904	4	US-08-961-527-238
12	40.8	6.6	84495	4	US-09-797-906-3
13	36.4	6.0	3487	4	US-09-453-702B-164
14	34.6	5.7	246240	2	US-08-724-394A-20
15	34.6	5.7	246240	2	US-08-724-394A-21
16	34.6	5.7	246240	2	US-08-724-394A-22
17	33.6	5.6	1245	4	US-09-266-910-1
18	33.6	5.6	1245	4	US-09-266-910-2
19	33.6	5.6	1711	1	US-08-568-147B-1
20	33.4	5.5	6669	3	US-09-212-971-5
21	33.4	5.5	6669	3	US-08-800-929A-5
22	33.4	5.5	6669	4	US-09-617-053A-5
23	33.4	5.5	162450	4	US-09-345-882-1
24	32.8	5.4	454	4	US-09-221-017B-179
25	32.8	5.4	889	4	US-09-513-783A-11
26	32.8	5.4	12571	4	US-09-322-478-20
27	32.6	5.4	87563	4	US-09-453-702B-57

28	32.2	5.3	3947	4	US-08-975-762-47	Sequence 47, App1
29	32.2	5.3	3947	4	US-09-295-028-47	Sequence 47, App1
30	32.2	5.3	3947	4	US-09-106-582-47	Sequence 47, App1
31	32	5.3	8920	2	US-08-446-855A-1	Sequence 1, App1
32	32	5.3	8920	2	US-09-150-741-1	Sequence 1, App1
33	32	5.3	9793	1	US-08-470-202-56	Sequence 56, App1
34	32	5.3	9793	1	US-08-471-770-56	Sequence 56, App1
35	32	5.3	9793	2	US-08-468-059-56	Sequence 56, App1
36	32	5.3	9793	4	US-09-109-916-56	Sequence 56, App1
37	31.4	5.2	1001	4	US-09-641-638-260	Sequence 260, App
38	31.4	5.2	2797	4	US-09-453-702B-244	Sequence 244, App
39	31.4	5.2	3393	4	US-09-104-324B-1	Sequence 1, App1
40	31.4	5.2	3393	4	US-09-162-713-1	Sequence 1, App1
41	31.4	5.2	3580	4	US-09-081-345-1	Sequence 1, App1
42	31.4	5.2	4291	2	US-08-417-210A-81	Sequence 81, App1
43	31.4	5.2	4517	4	US-09-140-804-9	Sequence 9, App1
44	31.4	5.2	9840	4	US-09-534-638-1	Sequence 1, App1
45	31.2	5.2	81001	4	US-09-750-580-1	Sequence 1, App1

ALIGNMENTS

RESULT 1
US-08-916-902A-2
; Sequence 2, Application US/08916902A
; Patent No. 5871930
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HIGH AFFINITY IMMUNOGLOBULIN E
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/916,902A
; FILING DATE: Herewith
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0371 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELETYPE:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1669 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAINOT04
; CLONE: 927955
; US-08-916-902A-2
Query Match 14.3% Score 86.4; DB 2; Length 1669;

Best Local Similarity 51.1%; Pred. No. 7e-15;
Matches 232; Conservative 0; Mismatches 216; Indels 6; Gaps 1;

```

QY 119 AAAAAATATTTGCTAGAAAAATGAAATCTTACGAGCTATACACCTGTTGGAATTA 178
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 325 AGAAGTTCTTGAAGGAGAACCCCAAGCTCTTGGGGTTGTGAGATTCGACTGCCCTGA 384
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 179 TGACCTTTCTTTGGAGTTATCTCTTTCACCTGTTAAAAACCATATCCAGGTTTC 238
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 385 TGAGCCTTACCATGGGAATTAACAATGATGTGATGGCATCTAATACCTTATGAGAATACC 444
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 239 CCTTTATATTTCTTACAGATATCCATCTGGGGCTCTGTTTGTTCATTATTCGGAG 298
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 445 CTATTTCCGGTATATCGGGGTACACAATTTGGGGGTACAGTAATGTTATTTATTCAGGAT 504
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 299 CCTTCTAATTTGAGTGAAGAAAAACACAGAAACTCTGATTAATATGAGCCGATTA 358
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 505 CCTTGCATATTCAGCAGAGAAATTAAGAACTACAAAGAGCCCTGGTCCGAGGATGCTAGGA 564
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 359 TGAATCTTCTTACGTGCGCTGAGAGCAATAGCTGAAATCTCTCCACATTTGGTTTCA 418
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 565 TGAATATACACAGCTCTGTACTGCTGCATCAGGAACTTAATACACATTTAGCTTG 624
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 419 TCCTAGAT-----CAAACTACATTTGTGTATTTCTCACCAAAATAGTCACTGAAG 472
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 625 CGTTTATTCATTCATACACCTTACTACTGATGCACTAGCAACTCAAAATATTTGCATG 684
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 473 CTGTACTGTCCTGTTCTTGGAAATTTGATTAATGATGACTTACAGCATTTATGAAT 532
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 685 GGACTATATGCCATCTTAATAGGCTGTGATGGCATGTGGTCCCTTAAAGTGTCTGAAT 744
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 533 TATTCATTTCTGCTGCTTCTCAATTTTGGGGTG 566
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 745 TCTGCATGTGCTGTCTCTCTCTGCTTGGATG 778
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 2

US-09-213-389-2
; Sequence 2, Application US/09213389
; Patent No. 5977072

GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
TITLE OF INVENTION: HIGH AFFINITY IMMUNOGLOBULIN E
TITLE OF INVENTION: RECEPTOR-LIKE PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Inocyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/213,389
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/916,902

FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0371 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166

TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1669 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAINOT04
; CLONE: 927955

US-09-213-389-2

Query Match 14.3%; Score 86.4; DB 2; Length 1669;
Best Local Similarity 51.1%; Pred. No. 7e-15;
Matches 232; Conservative 0; Mismatches 216; Indels 6; Gaps 1;

```

QY 119 AAAAAATATTTGCTAGAAAAATGAAATCTTACGAGCTATACACCTGTTGGAATTA 178
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 325 AGAAGTTCTTGAAGGAGAACCCCAAGCTCTTGGGGTTGTGAGATTCGACTGCCCTGA 384
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 179 TGACCTTTCTTTGGAGTTATCTCTTTCACCTGTTAAAAACCATATCCAGGTTTC 238
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 385 TGAGCCTTACCATGGGAATTAACAATGATGTGATGGCATCTAATACCTTATGAGAATACC 444
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 239 CCTTTATATTTCTTACAGATATCCATCTGGGGCTCTGTTTGTTCATTATTCGGAG 298
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 445 CTATTTCCGGTATATCGGGGTACACAATTTGGGGGTACAGTAATGTTATTTATTCAGGAT 504
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 299 CCTTCTAATTTGAGTGAAGAAAAACACAGAAACTCTGATTAATATGAGCCGATTA 358
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 505 CCTTGCATATTCAGCAGAGAAATTAAGAACTACAAAGAGCCCTGGTCCGAGGATGCTAGGA 564
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 359 TGAATCTTCTTACGTGCGCTGAGAGCAATAGCTGAAATCTCTCCACATTTGGTTTCA 418
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 565 TGAATATACACAGCTCTGTACTGCTGCATCAGGAACTTAATACACATTTAGCTTG 624
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 419 TCCTAGAT-----CAAACTACATTTGTGTATTTCTCACCAAAATAGTCACTGAAG 472
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 625 CGTTTATTCATTCATACACCTTACTACTGATGCACTAGCAACTCAAAATATTTGCATG 684
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 473 CTGTACTGTCCTGTTCTTGGAAATTTGATTAATGATGACTTACAGCATTTATGAAT 532
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 685 GGACTATATGCCATCTTAATAGGCTGTGATGGCATGTGGTCCCTTAAAGTGTCTGAAT 744
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 533 TATTCATTTCTGCTGCTTCTCAATTTTGGGGTG 566
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 745 TCTGCATGTGCTGTCTCTCTCTGCTTGGATG 778
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 3

US-08-318-492-3
; Sequence 3, Application US/08318492
; Patent No. 5552312

GENERAL INFORMATION:
APPLICANT: Lim, Bing
APPLICANT: Adra, Chaker N.
TITLE OF INVENTION: RECOMBINANT HTM4 GENE, PROTEIN AND
TITLE OF INVENTION: ASSAYS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02173

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:


```
APPLICATION NUMBER: US/08/318,492
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: BROOK, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: BI94-03
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1661 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 97..741
US-08-318-492-3
```

```
Query Match          10.5%; Score 63.2; DB 1; Length 1661;
Best Local Similarity 50.9%; Pred. No. 1.9e-08;
Matches 239; Conservative 0; Mismatches 213; Indels 18; Gaps 3;
```

```
QY 135 AAAAATGAATCTTGGGACTTCAGATCCGTTTGGAAATATGACCTTTCTTTGG 194
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 234 AAAATACAGTCTTGGGGCCATCCAGATCCGAATGCGAATATCTGCTTTGGG 293
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 195 AGT-----TATCTCTTTTCACTTGTAAACCATCCAAAGTTTCCCTTAT 245
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 294 TGTCTTTTGGGTTCTTGCAATACCATCCATCCACTTCCAAAGCAGCTTTTCTT 353
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 246 ATTCTTTAGATATCCATTCAGGGCTCTGTTTGTTCATTAATTCGAGCCTTCT 305
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 354 CTCTACACAGGCTACCCGATGGGGTGTGTTTCTGTAGTTCAGGAACCTTGTG 413
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 306 AATTGACGTAAAGAAAACACAGAACTCTGATTAATTTAGCCGAATTAATGATCT 365
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 414 TGTGTAGCAGGAGATAAACCCACAGAACATGATACAGAACGTTTGGAAATGACAT 473
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 366 TCTTAGGCGCTGAGACATATAGTGAATTCCTCCATTTGCTTCACTCTAG 425
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 474 TGCAGTGTACATTTGCACTAGTGGGAGCTGCTTTCTCTCAATAATATAGCAGTTAA 533
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 426 TCAAACTACATTTTGGTTAT-TCTACCAAAATAGTCAAGTAAAGCTGTACT---- 480
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 534 TATCCAGTATTAAGAGTTGTCACTTCATGAGAGTCAACCGACCTATGCAATTACAT 593
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 481 -----GTCCGTGTTTGGGATTTGATTTACATTGATGACTTTCAGCATTAATTAAT 536
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 594 GGGCTCAATATCAAAATGAGATGTGTCTACTGCTGATTCACCTTGTGGAATTAAG 653
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 537 CATTTCTGCGCTTCTCAATTTTGGGGGCCACCTGAGAGATTTGGATT 586
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 654 CGTAACATCTCTACCATAGCCATGTGTGCAATGCAAACTGCTTAAT 703
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 4
US-08-707-340-3
Sequence 3, Application US/08707340
Patent No. 5705615
GENERAL INFORMATION:
APPLICANT: Lim, Bing
APPLICANT: Adira, Chaker N.
APPLICANT: Leilas, Jean-Michel
TITLE OF INVENTION: RECOMBINANT HTM4 GENE, PROTEIN AND
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSER: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
```

```
CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/707,340
FILING DATE: 03-SEP-1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/318,492
FILING DATE: 06-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/675,648
FILING DATE: 03-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: BROOK, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: BI94-03A2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1661 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 97..738
US-08-707-340-3
```

```
Query Match          10.5%; Score 63.2; DB 1; Length 1661;
Best Local Similarity 50.9%; Pred. No. 1.9e-08;
Matches 239; Conservative 0; Mismatches 213; Indels 18; Gaps 3;
```

```
QY 135 AAAAATGAATCTTGGGACTTCAGATCCGTTTGGAAATATGACCTTTCTTTGG 194
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 234 AAAATACAGTCTTGGGGCCATCCAGATCCGAATGCGAATATCTGCTTTGGG 293
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 195 AGT-----TATCTCTTTTCACTTGTAAACCATATCCAAAGTTTCCCTTAT 245
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 294 TGTCTTTTGGGTTCTTGCAATACCATCCATCCACTTCCAAAGCAGCTTTTCTT 353
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 246 ATTCTTTAGATATCCATTCAGGGCTCTGTTTGTTCATTAATTCGAGCCTTCT 305
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 354 CTCTACACAGGCTACCCGATTTGGGGTGTGTTTCTGTAGTTCAGGAACCTTGTG 413
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 306 AATTGACGTAAAGAAAACACAGAACTCTGATTAATTTAGCCGAATTAATGATCT 365
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 414 TGTGTAGCAGGAGATAAACCCACAGAACATGATACAGAACGTTTGGAAATGACAT 473
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 366 TCTTAGGCGCTGAGACATATAGTGAATTCCTCCATTTGCTTCACTCTAG 425
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 474 TGCAGTGTACAAATTTGCACTAGTGGGAGCTGCTTTCTCTCAATAATATAGCAGTTAA 533
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 426 TCAAACTACATTTTGGTTAT-TCTACCAAAATAGTCAAGTAAAGCTGTACT---- 480
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 534 TATCCAGTATTAAGAGTTGTCACTTCATGAGAGTCAACCGACCTATGCAATTACAT 593
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 481 -----GTCCGTGTTTGGGATTTGATTTACATTGATGACTTTCAGCATTAATTAAT 536
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 594 GGGCTCAATATCAAAATGAGATGTGTCTACTGCTGATTCACCTTGTGGAATTAAG 653
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 537 CATTTCTGCGCTTCTCAATTTTGGGGGCCACCTGAGAGATTTGGATT 586
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 654 CGTAACATCTCTACCATAGCCATGTGTGCAATGCAAACTGCTTAAT 703
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```


[illegible]

EARLIER FILING DATE: 1997-06-13
 EARLIER APPLICATION NUMBER: 60/061,060
 EARLIER FILING DATE: 1997-10-02

Query Match 8.7%; Score 52.6; DB 4; Length 945;
 Best Local Similarity 51.2%; Pred. No. 1.4e-05;
 Matches 152; Conservative 0; Mismatches 139; Indels 6; Gaps 1;

QY 114 CTGCAAAATTTTGTCTGAGAAAATGAAATCTTAGGACATCATCTGCTTGG 173
 DB 147 CCGAAGAAACATCTACACCCAGAAATCAAGTTATGGACATCTCTGCTGG 206
 QY 174 AATTATGACCTTTCTTTGGAGTATCTTCCCTTACCTTTAAACCATAT---- 228
 DB 207 CATGATGATTTAGCTGGGAGATTTTGGACATCTTCTCTTCCCAATTTTAC 266
 QY 229 -CCAAGTTCCCTTATATTTCTTCAGATATCATTTGGGGCTCTTTTGTTCAT 287
 DB 267 CCAAGTACTTACACTGTGGAACCTGCTTACCATTCATAGGACCTTTTTTTTAT 326
 QY 288 TAATCTGAGCCTTCTTATTTGACATGAGAAAAGAAACACAGAACTGTATATAT 347
 DB 327 CATCTGTGCTCTCTATCAATCCACAGAGAAAGTTTACCAAGCTTTGGTGCATG 386
 QY 348 GAGCCGATATGATATCTTCTTAGTGCCCTGAGAGCAATGCTGGAATCTCTCT 404
 DB 387 CACCTGCTGTGGAAGATTTCTGAGTGTCTGTGCCCTGGTGGTTTCAATTATCT 443

RESULT 7
 US-07-869-933-22
 Sequence 22, Application US/07869933
 Patent No. 5770396

GENERAL INFORMATION:
 APPLICANT: KINET, Jean-Pierre
 TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF
 TITLE OF INVENTION: THE HUMAN B SUBUNIT OF THE HIGH AFFINITY RECEPTOR FOR
 NUMBER OF SEQUENCES: 34
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Foley & Lardner
 STREET: 1800 Diagonal Road, Suite 500
 CITY: Alexandria
 STATE: VA
 COUNTRY: USA
 ZIP: 22313-0299

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/869,933
 FILING DATE: 19920416

CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: BENT, Stephen A.
 REGISTRATION NUMBER: 29,768
 REFERENCE/DOCKET NUMBER: 40399/154 NIH
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703)836-9300
 TELEFAX: (703)836-4109
 TELEX: 899149
 INFORMATION FOR SEQ ID NO: 22:
 SEQUENCE CHARACTERISTICS:

LENGTH: 2545 base pairs
 TYPE: NUCLEIC ACID
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 46..786

FEATURE:
 NAME/KEY: sig_peptide
 LOCATION: 46..54
 FEATURE:
 NAME/KEY: mat_peptide
 LOCATION: 55..786
 US-07-869-933-22

Query Match 8.2%; Score 49.2; DB 1; Length 2545;
 Best Local Similarity 52.3%; Pred. No. 0.00017;
 Matches 158; Conservative 0; Mismatches 138; Indels 6; Gaps 2;

QY 101 CAACCAAGCCCTTGCAAAATTAATTTGCTAGAAAATGAAATCTTAGGACATATCC 160
 DB 188 CACCAGCAGACATGACACTGATTTTGAAGAAAGTTGGAGTTCCTGGGGTACCC 247
 QY 161 AGATCTGTTGGAAATTAATACCTTTTCTTTGGAGTATCTTCCTTTACCTT----- 215
 DB 248 AAGTTCTGGTTGGTTGATATGCTTGTGTTGGAACAGTTGCTGCTCCACACTCCAGA 307
 QY 216 GTAAACCATATCCAAAGTTTCCCTTTA-TATTTCTTCAGATATCATCTGGGGCT 274
 DB 308 CTTCAGACTTTGACGAGAAATCTTTTATTAATATAGCAGCTTACCATTTCTGGGGTG 367
 QY 275 CTGTTTGTTCATTAATTTCTGAGCCTTCTCAATTCAGTGAAGAAAGAAACACAGAAA 334
 DB 368 CAGTGTGTTGTTTGTCTGATTTTGTTCATTTATGTCGAAAGAAACACACTGT 427
 QY 335 CTCTGATATATTTGAGCCGAAATATGATCTTCTAGTCCCTGAGAGCAATAGCTGGAA 394
 DB 428 ATCTGTGAGAGCAGCCTGGGAGCAAAATTTCTCAGCAGCATCGCTCGAGCTTGGGA 487
 QY 395 TC 396
 DB 488 TC 489

RESULT 8
 US-09-103-663-22
 Sequence 22, Application US/09103663D
 Patent No. 6171803

GENERAL INFORMATION:
 APPLICANT: Kinet et al.
 TITLE OF INVENTION: Isolation, characterization, and use of the human beta
 TITLE OF INVENTION: subunit of the high affinity receptor for
 FILE REFERENCE: 50490
 CURRENT APPLICATION NUMBER: US/09/103,663D
 CURRENT FILING DATE: 1998-06-23
 EARLIER APPLICATION NUMBER: 07/869,933
 EARLIER FILING DATE: 1992-04-16
 NUMBER OF SEQ ID NOS: 35
 SOFTWARE: Patent Ver. 2.1

SEQ ID NO 22
 LENGTH: 2545
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: sig_peptide
 LOCATION: (46)..(54)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (46)..(786)
 US-09-103-663-22

Query Match 8.2%; Score 49.2; DB 4; Length 2545;
 Best Local Similarity 52.3%; Pred. No. 0.00017;
 Matches 158; Conservative 0; Mismatches 138; Indels 6; Gaps 2;

QY 101 CAACCAAGCCCTTGCAAAATTAATTTGCTAGAAAATGAAATCTTAGGACATATCC 160
 DB 188 CACCAGCAGACATGACACTGATTTTGAAGAAAGTTGGAGTTCCTGGGGTACCC 247

Query Match	8.0%	Score 48	DB 48	Length 1308
Best Local Similarity	50.7%	Pred. No. 0.00029		
Matches 145	Conservative 0	Mismatches 135	Indels 6	Gaps 1
QY	137	AAATGAAAATCTTAAAGGACTATCCAGATCCGTTTGGATTAATGACCTTTCTTTGGAG	196	
Db	195	AGATCAAAAGTATGGCGGCATCCAGATCATGTGCTGTGATGGTGTGATCGTGGAA	254	
QY	197	TTATCTT-----CGTTTCACTCTGTTAAACCATATCCAGAGTTTCCCTTAATATTC	250	
Db	255	TCATTTTGGCATCTGTTTCCCTCCCAATCTACACTTTACCTCAAGTGTTTTCCACCTGGT	314	
QY	251	TTTCAGGATATCCATCTTGGGGCTCTGTTTGTTCATTAAATTCGAGGCTTCCAAATG	310	
Db	315	AATCTGGCTACCCATTTGTAGAGCTGTTTGTTCATCTCGAATATCTGCTAATG	374	
QY	311	CAGTGAAGAAAACCCACAGAACTCTGATTAATATGCGCGAATTAATCTTCTTA	370	
Db	375	TCACAGAGAAAAGATGACTAAGCCTTTGGTTCCACAGACGCTACGCCCTGACATCTTA	434	
QY	371	GTCGCCCTGAGACCAATAGCTGGAATCATCTTCCCTCACATTTGGTTTT	416	
Db	435	GTCGCTCTCTGCTCTTACAGGCAATCGCTATTTCTCTGTCAAGTTT	480	

RESULT 10
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:

Query Match	7.1%	Score 42.6	DB 1	Length 7218
Best Local Similarity	13.5%	Pred. 0.017		
Matches 75	Conservative 213	Mismatches 267	Indels 0	Gaps 0
QY 6	TTCAAGCAGCCGACAGCTCCGGTGTTCCTGGATTTTCCTCCGAAATCAGTCTTCAGA	65		
DB 875	TTAAGCATTTTCATACACAGCAGGTAGTTTACCACCATTTTCAGTAATAGGAATFAMA	934		
QY 66	ATATGAGTCCACAGACATTTAGCCACGACCTTTTCAACTCAAGCCCTTGCAAAATTT	125		
DB 935	TTCAAAATATTTATTAAGGGTTTACGTTGAATGTCATCGAGCGGCTACTACTTA	994		
QY 126	ATTTGCTGAAATAAGAAATCTTAGGAGCTATCCAGATTCGCTTTGGAAATTAATACCTT	185		
DB 995	ATTTTCTCTTGCTTGTCGCATACGCTCACAGATTAATTTCCGACTTGCGTGCAGSTCGAG	1054		
QY 186	TTCTTTTGGAGTTACTTCTCTTTCCACCTGTGTAAACCATATCCAGAGTTTCCCTTAT	245		
DB 1055	GGACGCTTCGATTT	1114		
QY 246	ATTTCCTTCAGAGATATCATCTCGGGCTCTGTTTGTCTAATTAATCTCGAGCCTTCT	305		
DB 1115	TT	1174		
QY 306	AATTCGAGTGAAGAAAGAAACACACAGAACTGTATATATGAGGCCAATAATGAATCT	365		
DB 1175	TT	1234		
QY 366	TCTATGATGCCCTGAGAGCAATAGCTGGAATATCTCTCCACATTTGGTTTCAATCTGA	425		

Db 1235 YY 1294
QY 426 TCAGAACTACATTTGGTATATCTCACCAAAATAGTCACTGTAAGGCTTTACTGCTT 485
Db 1295 YY 1354
QY 486 GTTCCTTGGAAATTTGATATACATTTGATGACTTTTCACATATTTGAATTTACTTCTCT 545
Db 1355 YY 1414
QY 546 GCCTTTCTCATTTT 560
Db 1415 YYYYYYYYYYYYYY 1429

RESULT 11
US-08-961-527-238/c
; Sequence 238, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 238:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 904 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-961-527-238

Query Match 6.8%; Score 40.8; DB 4; Length 904;
Best Local Similarity 54.7%; Pred. No. 0.026;
Matches 81; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 423 AGATCAAACTACATTTGGTATATCTCACCAAAATAGTCACTGTAAGGCTTTACTGT 482
Db 610 AGTTAAAGAGATCTTGATACACGAGAAAAAGACACATTTATGACATGACAGC 551
QY 483 CCTGTTCTTGGCAATTTGATATACATTTGATGACTTTTCACATATTTGAATTTACTTTC 542
Db 550 CCTTTAAGGGATATGACATTTGATTCATTTGATGCTTTCAAGATGTCATTC 491
QY 543 TCGCCTTTCTCAATTTTGGGTCAC 570
Db 490 ACCTGCTCTATACTTTGGGAGCCAC 463

RESULT 12
US-09-797-906-3/c
; Sequence 3, Application US/09797906
; Patent No. 6329188
; GENERAL INFORMATION:
; APPLICANT: Zianghe YAN, Karen A. KETCHUM, Valentina DIFRANCESCO, Ellen M. BEASLEY
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: C1001151CIP
; CURRENT APPLICATION NUMBER: US/09/797,906
; CURRENT FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 84495
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(84495)
; OTHER INFORMATION: n = A,T,C or G
US-09-797-906-3

Query Match 6.6%; Score 40; DB 4; Length 84495;
Best Local Similarity 44.9%; Pred. No. 0.21;
Matches 151; Conservative 0; Mismatches 185; Indels 0; Gaps 0;

QY 108 AAGCCCTTGCAAAATTTATTTGCTAGAAAAATGAAAATCTTAGGACTATCCAGATCCT 167
Db 66394 AAGACTATTTCCATTAACGCCATTAACATTAATTAATTTCTAAACAAAATTAATCTT 66335
QY 168 GTTGAATTAAGACCTTTCTTTGAGATATCTTCTTTCACCTTTGTAACCA 227
Db 66334 TTAATTAATAATTAACCATTTCTTAATTAATTTCTTCAATTTGTATATATATATAT 66275
QY 228 TCCAGGTTTCCCTTATATTTCTTTCAGATATCCATTCGTGGGCTGTTTGTTCAT 287
Db 66274 TACCGTGTGTAAGTTTATTAACAGAAAGATATCTGATGATTTTATCTTTGAGTC 66215
QY 288 TAATTCGAGCCCTTCTTAATTCAGTGAAGAAAACCAAGAACTCTGATAATAT 347
Db 66214 ACTTCTAATTCGTGCCACCTTTTGACAGAAAATAGCTCATATTTCTTTATATAC 66155
QY 348 GAGCCCAATTAATTAATCTTCTAGTGCCTGAGAGCAATAGCTGGAATATCTCTCAC 407
Db 66154 CTTATACAGAAAGGTCCTCAGATTTATTAAGAAAGACTAAATGAAGAACTAAATCTTT 66095
QY 408 ATTTGTTTCATCTGATCAAAACTACATTTGTGG 443
Db 66094 TTTTTTTTTTTTTTGAGATGAGACTCTCCTGTTG 66059

RESULT 13
US-09-453-702B-164
; Sequence 164, Application US/09453702B
; Patent No. 6365723
; GENERAL INFORMATION:
; APPLICANT: Blatner, Frederick R.
; Perna, Nicole T.
; Welch, Rod
; TITLE OF INVENTION: No. 6365723el Sequences of E. coli 0157
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Plinkney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US

```

;
; ZIP: 53701-2113
;
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: Diskette, 3.50 Inch, 1.44Mb storage
;
; COMPUTER: IBM PC compatible
;
; OPERATING SYSTEM: PC-DOS/MS-DOS
;
; SOFTWARE: Word Perfect 8.0
;
; CURRENT APPLICATION DATA:
;
; APPLICATION NUMBER: US/09/453,702B
;
; FILING DATE: 03-Dec-1999
;
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
;
; APPLICATION NUMBER: 60/110,955
;
; FILING DATE: 04-Dec-1998
;
; ATTORNEY/AGENT INFORMATION:
;
; NAME: Seay, Nicholas J.
;
; REGISTRATION NUMBER: 27386
;
; REFERENCE/DOCKET NUMBER: 960296,95017
;
; TELECOMMUNICATION INFORMATION:
;
; TELEPHONE: (608) 251-5000
;
; TELEFAX: (608) 251-9166
;
; INFORMATION FOR SEQ ID NO: 164:
;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 3487
;
; TYPE: nucleic acid
;
; STRANDEDNESS: double
;
; TOPOLOGY: linear
;
; MOLECULE TYPE: DNA (genomic)
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 164:
;
US-09-453-702B-164

```

```

Query Match
Best Local Similarity 51.9%; Score 36.4; DB 4; Length 3487;
Matches 82; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

```

```

QY 255 AGCATATCCATTCTGGGCTCTGTTTGTTCATTAATTCGAGCCTCCTAAATTCGAGT 314
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2325 ACGAAATTCATGTTGTCGTGATGCGATAAATTAATTCGAGTGTGCTGGTTCG 2384
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 315 GAAAGAAACACAGAACTCGAATATATGAGCGCAATATGAAATCTTCTTAATGC 374
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2385 CAGTAATTAAGATTTGCAATTCATTCACAAATTAATTTGGGTTAATGATCATTTTATCGG 2444
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 375 CCGAGAGCAATAGCTGATCATTCCTCCACATTTG 412
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2445 CCCAATACCATTCGTGCGATCGTTCGTCGATCATTTG 2482
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 14
US-08-724-394A-20
; Sequence 20, Application US/08724394A
; Patent No. 5872237
;
; GENERAL INFORMATION:
;
; APPLICANT: Feder, John N.
;
; APPLICANT: Kironmal, Gregory S.
;
; APPLICANT: Laufer, Peter M.
;
; APPLICANT: Ruddy, David A.
;
; APPLICANT: Thomas, Winston
;
; APPLICANT: Tsuchihashi, Zenta
;
; APPLICANT: Wolff, Roger K.
;
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
;
; NUMBER OF SEQUENCES: 31
;
; CORRESPONDENCE ADDRESS:
;
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
;
; STREET: Two Embarcadero Center, 8th Floor
;
; CITY: San Francisco
;
; STATE: CA
;
; COUNTRY: USA
;
; ZIP: 94111-3834
;
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: Floppy disk
;
; COMPUTER: IBM PC compatible
;
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

;
; SOFTWARE: Patentin Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
;
; APPLICATION NUMBER: US/08/724,394A
;
; FILING DATE: 01-OCT-1996
;
; CLASSIFICATION: 536
;
; ATTORNEY/AGENT INFORMATION:
;
; NAME: Filts, Renee A.
;
; REGISTRATION NUMBER: 35,136
;
; REFERENCE/DOCKET NUMBER: 017957-000100
;
; TELECOMMUNICATION INFORMATION:
;
; TELEPHONE: 415-576-0200
;
; TELEFAX: 415-576-0300
;
; INFORMATION FOR SEQ ID NO: 20:
;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 246240 base pairs
;
; TYPE: nucleic acid
;
; STRANDEDNESS: not relevant
;
; TOPOLOGY: not relevant
;
; MOLECULE TYPE: CDNA
;
; FEATURE:
;
; NAME/KEY: misc.feature
;
; LOCATION: 1..246240
;
; OTHER INFORMATION: /note= "H1A-H.CONTIG"
;
US-08-724-394A-20

```

```

Query Match
Best Local Similarity 55.4%; Score 34.6; DB 2; Length 246240;
Matches 67; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

```

```

QY 449 CTCACCAAAATAGTCACTGTAAGGCTGTCCTGCTCTGTTCTGGGAATTTTGAATTCAT 508
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 36690 CTCCTCGAGATCTGCTGTAATTAATTCATTTATTCGTGTGTGTAATTAACCAACATG 36749
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 509 TGAATGACTTTCAGCATTAATGAATTAATTCATTTCTGCTTCTTCATTTTGGGGTCC 568
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 36750 TGAATGCTTGAACACACAAATTAATTTATTTATTTACAAATTCCTAAGCTCAGAGTCC 36809
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 569 A 569
DB 36810 A 36810

```

```

RESULT 15
US-08-724-394A-21
; Sequence 21, Application US/08724394A
; Patent No. 5872237
;
; GENERAL INFORMATION:
;
; APPLICANT: Feder, John N.
;
; APPLICANT: Kironmal, Gregory S.
;
; APPLICANT: Laufer, Peter M.
;
; APPLICANT: Ruddy, David A.
;
; APPLICANT: Thomas, Winston
;
; APPLICANT: Tsuchihashi, Zenta
;
; APPLICANT: Wolff, Roger K.
;
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
;
; NUMBER OF SEQUENCES: 31
;
; CORRESPONDENCE ADDRESS:
;
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
;
; STREET: Two Embarcadero Center, 8th Floor
;
; CITY: San Francisco
;
; STATE: CA
;
; COUNTRY: USA
;
; ZIP: 94111-3834
;
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: Floppy disk
;
; COMPUTER: IBM PC compatible
;
; OPERATING SYSTEM: PC-DOS/MS-DOS
;
; SOFTWARE: Patentin Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
;
; APPLICATION NUMBER: US/08/724,394A
;
; FILING DATE: 01-OCT-1996
;
; CLASSIFICATION: 536

```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 22, 2003, 10:40:16 : Search time 52 seconds
(without alignments)
6512.857 Million cell updates/sec

Title: US-09-735-712-1

Perfect score: 603

Sequence: 1 atgatttcacagcaccgcacac.....attgtgaacatgttgttga 603

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 442118 seqs, 280819700 residues

Total number of hits satisfying chosen parameters: 884236

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications: NA.*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	603	100.0	603	10	US-09-735-712-1
2	603	100.0	603	10	US-09-735-712-9
3	599.8	99.5	723	9	US-09-924-340-33
4	599.8	99.5	723	9	US-09-992-600A-33
5	598.2	99.2	747	10	US-09-731-872-17
6	598.2	99.2	760	10	US-09-821-821-1
7	587.8	97.5	689	9	US-09-949-842-10
8	347.4	57.6	450	10	US-09-735-712-7
9	341.8	56.7	417	10	US-09-735-712-5
10	340.2	56.4	382	10	US-09-964-824A-207
11	246	40.8	246	10	US-09-735-712-3
12	158.4	26.3	468	10	US-09-731-872-21
13	86.4	14.3	2438	9	US-09-822-846-227
14	62.6	10.4	1563	10	US-09-765-205-21
15	55	9.1	470	9	US-09-796-692-6874
16	55	9.1	610	9	US-09-796-692-2689
17	55	9.1	1597	10	US-09-954-456-734
18	54.6	9.1	615	9	US-09-796-692-4034
19	52.6	8.7	1728	10	US-09-822-849A-23

20	52.6	8.7	1743	10	US-09-925-302-314	Sequence 314, App
21	52.4	8.7	573	9	US-09-796-692-4497	Sequence 4497, Ap
22	52	8.6	405	9	US-09-796-692-2750	Sequence 2750, Ap
23	52	8.6	1317	9	US-09-981-953-81	Sequence 81, Appl
24	51.8	8.6	470	9	US-09-796-692-6989	Sequence 6989, Ap
25	51.6	8.6	1291	10	US-09-739-254-17	Sequence 17, Appl
26	51.6	8.6	1291	10	US-09-904-615-17	Sequence 17, Appl
27	49.4	8.2	982	10	US-09-821-821-3	Sequence 3, Appl
28	48.6	8.1	333	9	US-09-796-692-2788	Sequence 2788, Ap
29	48.4	8.0	484	9	US-09-796-692-7	Sequence 7, Appl
30	48.4	8.0	1292	10	US-09-739-254-61	Sequence 311, Ap
31	48.4	8.0	1292	10	US-09-904-615-61	Sequence 61, Appl
32	48.4	8.0	1292	10	US-09-904-615-61	Sequence 61, Appl
33	48.2	8.0	356	9	US-09-796-692-2735	Sequence 2735, Ap
34	48	8.0	571	9	US-09-796-692-363	Sequence 363, App
35	48	8.0	571	9	US-09-796-692-4970	Sequence 4970, Ap
36	48	8.0	1308	9	US-10-152-661-581	Sequence 581, App
37	47.6	7.9	600	9	US-09-796-692-6877	Sequence 6877, App
38	43.6	7.2	617	9	US-09-796-692-9117	Sequence 9117, Ap
39	42.6	7.1	553	9	US-09-796-692-171	Sequence 171, App
40	42.6	7.1	553	9	US-09-796-692-9488	Sequence 9488, App
41	40.2	6.7	348	9	US-09-796-692-3476	Sequence 3476, Ap
42	39.6	6.6	502	9	US-09-796-692-4316	Sequence 4316, Ap
43	39.6	6.6	513509	9	US-09-754-853A-4	Sequence 4, Appl
44	38.6	6.4	352	10	US-09-783-590-4341	Sequence 4341, Ap
45	38	6.3	350	10	US-09-770-791-697	Sequence 697, App

ALIGNMENTS

RESULT 1
US-09-735-712-1
; Sequence 1, Application US/09735712
; Patent No. US20020045743A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Made
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0109-USA
; CURRENT APPLICATION NUMBER: US/09/735,712
; CURRENT FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 60/171,567
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 603
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-735-712-1

Query Match 100.0%; Score 603; DB 10; Length 603;

Best local similarity 100.0%; Pred. No. 3.2e+143; Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGATTCAAGCACCACAGCTCCGGTGTTCGTGATTTCTCCAGAAATCACTGCT	60
DB	1	ATGATTCAAGCACCACAGCTCCGGTGTTCGTGATTTCTCCAGAAATCACTGCT	60
QY	61	TCGAAATGAGTCCACAGACTTTCAGCCACACCTTTCACTCAAAAGCCCTTGCAA	120
DB	61	TCGAAATGAGTCCACAGACTTTCAGCCACACCTTTCACTCAAAAGCCCTTGCAA	120
QY	121	AAATTTTCTGAGAAATGAATCTTAGGACTATCCAGATCCCTGTTGGAATTATG	180
DB	121	AAATTTTCTGAGAAATGAATCTTAGGACTATCCAGATCCCTGTTGGAATTATG	180
QY	181	ACCTTTTCTTGGAGTATCTTCTTTTCACTTGTAAACATATCCAGGTTTCCC	240
DB	181	ACCTTTTCTTGGAGTATCTTCTTTTCACTTGTAAACATATCCAGGTTTCCC	240

QY 241 TTTAATTTCTTTCAGAGATATCCATTCTGGGCTCTGTTTGTTCATTAATTCGAGCC 300
DB 241 TTTAATTTCTTTCAGAGATATCCATTCTGGGCTCTGTTTGTTCATTAATTCGAGCC 300
QY 301 TTCCTAATTCGAGTGAAGAAAACCCAGAACTCTATATATATGAGCCGAATATG 360
DB 301 TTCCTAATTCGAGTGAAGAAAACCCAGAACTCTATATATATGAGCCGAATATG 360
QY 361 AATCTTCTAGTCCCTGAGAGCAATAGCTGAATCATTCCTCCATTTGGTTTCATC 420
DB 361 AATCTTCTAGTCCCTGAGAGCAATAGCTGAATCATTCCTCCATTTGGTTTCATC 420
QY 421 CTAGATCAAACTACATTTGTGTATTCCTACCAAAATAGTCAGTGAAGCTGTACT 480
DB 421 CTAGATCAAACTACATTTGTGTATTCCTACCAAAATAGTCAGTGAAGCTGTACT 480
QY 481 GTCCCTGTCTTGGGAATTTTGATTAATGATGACTTTAGCATTTATTAATTCATT 540
DB 481 GTCCCTGTCTTGGGAATTTTGATTAATGATGACTTTAGCATTTATTAATTCATT 540
QY 541 TCTCTGCTTTTCATTTTGGGGTCCACTCAGAGATTGTGATGTGAACATGTGT 600
DB 541 TCTCTGCTTTTCATTTTGGGGTCCACTCAGAGATTGTGATGTGAACATGTGT 600
QY 601 TGA 603
DB 601 TGA 603

RESULT 2

US-09-735-712-9
; Sequence 9, Application US/09735712
; Patent No. US20020045743A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; TITLE OF INVENTION: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20020045743A1 Human Membrane Proteins and
; FILE REFERENCE: Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0109-USA
; CURRENT APPLICATION NUMBER: US/09/735,712
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 60/171,567
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 676
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-735-712-9

Query Match 100.0%; Score 603; DB 10: Length 676;
Best Local Similarity 100.0%; Pred. No. 3.3e-143;
Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGATTCAGACGACGACACAGTCGCGTGTTCGTGTAATTCCTCCAGAAATCACTGCT 60
DB 33 ATGATTCAGACGACGACACAGTCGCGTGTTCGTGTAATTCCTCCAGAAATCACTGCT 92
QY 61 TCAGAAATAGATCCACAGACTTTCAGCCAGACCTTTTCAACTCAAAAGCCCTTGCAG 120
DB 93 TCAGAAATAGATCCACAGACTTTCAGCCAGACCTTTTCAACTCAAAAGCCCTTGCAG 152
QY 121 AATATATTCGTAGAAAATGAAAATCTAGGAGATCCAGATCCATCTGTTGGAATATG 180
DB 153 AATATATTCGTAGAAAATGAAAATCTAGGAGATCCAGATCCATCTGTTGGAATATG 212
QY 181 ACCTTTCTTTGAGATATCTTCTTCCATTTGTTAAACCAATATCAAGGTTTCC 240
DB 213 ACCTTTCTTTGAGATATCTTCTTCCATTTGTTAAACCAATATCAAGGTTTCC 272
QY 241 TTTAATTTCTTTCAGAGATATCCATTCTGGGCTCTGTTTGTTCATTAATTCGAGCC 300

DB 273 TTTAATTTCTTTCAGAGATATCCATTCTGGGCTCTGTTTGTTCATTAATTCGAGCC 332
QY 301 TTCCTAATTCGAGTGAAGAAAACCCAGAACTCTATATATGAGCCGAATATG 360
DB 333 TTCCTAATTCGAGTGAAGAAAACCCAGAACTCTATATATGAGCCGAATATG 392
QY 361 AATCTTCTAGTCCCTGAGAGCAATAGCTGAATCATTCCTCCATTTGGTTTCATC 420
DB 393 AATCTTCTAGTCCCTGAGAGCAATAGCTGAATCATTCCTCCATTTGGTTTCATC 452
QY 421 CTAGATCAAACTACATTTGTGTATTCCTACCAAAATAGTCAGTGAAGCTGTACT 480
DB 453 CTAGATCAAACTACATTTGTGTATTCCTACCAAAATAGTCAGTGAAGCTGTACT 512
QY 481 GTCCCTGTCTTGGGAATTTTGATTAATGATGACTTTAGCATTTATTAATTCATT 540
DB 513 GTCCCTGTCTTGGGAATTTTGATTAATGATGACTTTAGCATTTATTAATTCATT 572
QY 541 TCTCTGCTTTTCATTTTGGGGTCCACTCAGAGATTGTGATGTGAACATGTGT 600
DB 573 TCTCTGCTTTTCATTTTGGGGTCCACTCAGAGATTGTGATGTGAACATGTGT 632
QY 601 TGA 603
DB 601 TGA 603
QY 633 TGA 635
DB 633 TGA 635

RESULT 3

US-09-924-340-33
; Sequence 33, Application US/09924340
; Publication No. US20030027248A1
; GENERAL INFORMATION:
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US2.REG
; CURRENT APPLICATION NUMBER: US/09/924,340
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: Jpatent
; SEQ ID NO 33
; LENGTH: 723
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: 1..72
; NAME/KEY: CDS
; LOCATION: 73..672
; NAME/KEY: 3'UTR
; LOCATION: 673..723
; NAME/KEY: polyA_signal
; LOCATION: 689..694
; NAME/KEY: polyA_site
; LOCATION: 708..723
US-09-924-340-33

Query Match 99.5%; Score 599.8; DB 9: Length 723;
Best Local Similarity 99.7%; Pred. No. 2.2e-142;
Matches 601; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGATTCAGACGACGACACAGTCGCGTGTTCGTGTAATTCCTCCAGAAATCACTGCT 60
DB 73 ATGATTCAGACGACGACACAGTCGCGTGTTCGTGTAATTCCTCCAGAAATCACTGCT 132

```

: FEATURE:
: NAME/KEY: CDS
: LOCATION: 73..672
: FEATURE:
: NAME/KEY: 3'UTR
: LOCATION: 673..723
: FEATURE:
: NAME/KEY: polyA_signal
: LOCATION: 689..694
: FEATURE:
: NAME/KEY: polyA_site
: LOCATION: 708..723
US-09-992-600A-33

Query Match          99.5%; Score 599.8; DB 9; Length 723;
Best Local Similarity 99.7%; Pred. No. 2.2e-142;
Matches 601; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY      1  ATGAGATTCAGACCGCACACAGTCGGGTCTTGATTTCTCCAGAAATCACTGCT 60
Db       73  ATGAGATTCAGACCGCACACAGTCGGGTCTTGATTTCTCCAGAAATCACTGCT 132
OY      61  TCAGAAATFAGACTCCACAGAACTTTCAGCCACGACCTTTTCACTCAAGCCCTTGC 120
Db      133  TCAGAAATFAGACTCCACAGAACTTTCAGCCACGACCTTTTCACTCAAGCCCTTGC 192
OY     121  AAATATTTGCTAGAAAAATGAAATCTTAGGACATCCAAATCCTGTTGGAAATTAG 180
Db     193  AAATATTTGCTAGAAAAATGAAATCTTAGGACATCCAAATCCTGTTGGAAATTAG 252
OY     181  ACCTTTCTTTTGGAGTTATCTCTTTTACCTTGTAAACCAATFCCAAGTTTCCG 240
Db     253  ACCTTTCTTTTGGAGTTATCTCTTTTACCTTGTAAACCAATFCCAAGTTTCCG 312
OY     241  TTTATATTTCTTTCAGGATTCATTCGCGGCTCTGTTTGTCAATTAATCTGGAGCC 300
Db     313  TTTATATTTCTTTCAGGATTCATTCGCGGCTCTGTTTGTCAATTAATCTGGAGCC 372
OY     301  TTCCTAATTTGCAAGTAAAGAAAAACACAGAAACTCTGATTAATATTGAGCCGAATAATG 360
Db     373  TTCCTAATTTGCAAGTAAAGAAAAACACAGAAACTCTGATTAATATTGAGCCGAATAATG 432
OY     361  AATCTCTTTAGTGGCCCTGAGAGCAATAGCTGGAATCATTTCTCTCAATTTGGTTTCAATC 420
Db     433  AATTTCTTTAGTGGCCCTGAGAGCAATAGCTGGAATCATTTCTCTCAATTTGGTTTCAATC 492
OY     421  CTAGATCAAAACTTACATTTGGTATTCCTACCAAAATAGCAGNGTAAGCTGTACT 480
Db     493  CTAGATCAAAACTTACATTTGGTATTCCTACCAAAATAGCAGNGTAAGCTGTACT 552
OY     481  GTCCGTCTCTGGGAATTTTGATTCATTTGATGATGACTTTCAGCAATTAATTAATTCATT 540
Db     553  GTCCGTCTCTGGGAATTTTGATTCATTTGATGATGACTTTCAGCAATTAATTAATTCATT 612
OY     541  TCTCTGCGCTTCTCAATTTTGGGGTGCACCTCAGAGGAAATTTGATTTGGAACAATGTTGT 600
Db     613  TCTCTGCGCTTCTCAATTTTGGGGTGCACCTCAGAGGAAATTTGATTTGGAACAATGTTGT 672
OY      601  TGA 603
Db      673  TGA 675

RESULT 5
US-09-731-872-17
: Sequence 17, Application US/09731872
: Patent No. US20020102604A1
: GENERAL INFORMATION:
: APPLICANT: Dumas Milne Edwards, Jean Baptiste
: APPLICANT: Bouquelieret, Lydie
: TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
: FILE REFERENCE: 78.US3.REG

```

```
; CURRENT APPLICATION NUMBER: US/09/731,872
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 17
; LENGTH: 747
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 73..672
; NAME/KEY: sig.peptide
; LOCATION: 73..132
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 5.21332530399231
; OTHER INFORMATION: seq SPVFLVPEPEITA/SE
US-09-731-872-17
```

```
Query Match          99.2%; Score 598.2; DB 10; Length 747;
Best Local Similarity 99.5%; Pred. No. 5,6e-142;
Matches 600; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 ATGATTTCAAGCAGCCGACAGAGTCGGGTCTGTATTTCTCCAGAAATCACTGCT 60
   |||||||
Db 73 ATGATTTCAAGCAGCCGACAGAGTCGGGTCTGTATTTCTCCAGAAATCACTGCT 132
   |||||||
QY 61 TCAGAAATAGTTCACAGAACTTTCAGCAGACCTTTTCAACTCAAAAGCCCTTGCAA 120
   |||||||
Db 133 TCAGAAATAGTTCACAGAACTTTCAGCAGACCTTTTCAACTCAAAAGCCCTTGCAA 192
   |||||||
QY 121 AATTTTTCGTAGAAATGAAATCTTAGGACTATCCAGATCTCTGTTGGAATTATG 180
   |||||||
Db 193 AATTTTTCGTAGAAATGAAATCTTAGGACTATCCAGATCTCTGTTGGAATTATG 252
   |||||||
QY 181 ACCTTTCTTTGGAGATATCTCTTTCACCTGTAAACCATATCCAGAGTTTCC 240
   |||||||
Db 253 ACCTTTCTTTGGAGATATCTCTTTCACCTGTAAACCATATCCAGAGTTTCC 312
   |||||||
QY 241 TTTTATTTCTTTCAGGATATCCATCTGGGGCTCTGTTTGTCTAATTAATTCGAGGC 300
   |||||||
Db 313 TTTTATTTCTTTCAGGATATCCATCTGGGGCTCTGTTTGTCTAATTAATTCGAGGC 372
   |||||||
Y 301 TTCCTAATTCAGTGAAGAAAGAAACACAGAAACCTGTATATATTTGAGCCGAATATG 360
   |||||||
Db 373 TTCCTAATTCAGTGAAGAAAGAAACACAGAAACCTGTATATATTTGAGCCGAATATG 432
   |||||||
QY 361 AATCTTCTTAGTCCCTGAGAGCAATAGCTGGAATCATCTCTCCACATTTGGTTTATC 420
   |||||||
Db 433 AATTTTCTTAGTCCCTGAGAGCAATAGCTGGAATCATCTCTCCACATTTGGTTTATC 492
   |||||||
QY 421 CTAGATCAAAACTACTATTGTGTATTTCTACCAAAATATGTCAGTGAAGCTGTACT 480
   |||||||
Db 493 CTAGATCAAAACTACTATTGTGTATTTCTACCAAAATATGTCAGTGAAGCTGTACT 552
   |||||||
QY 481 GTCCTGTTCTTGGGAATTTGATTCATATGATGACTTTCAGATTAATTAATTCAT 540
   |||||||
Db 553 GTCCTGTTCTTGGGAATTTGATTCATATGATGACTTTCAGATTAATTAATTCAT 612
   |||||||
QY 541 TCTTGCCTTCTTCAATTTTGGGGTCCACTCAGAGATTTGATTTGAACAATGTTGT 600
   |||||||
Db 613 TCTTGCCTTCTTCAATTTTGGGGTCCACTCAGAGATTTGATTTGAACAATGTTGT 672
   |||||||
QY 601 TGA 603
   |||
Db 673 TGA 675
```

RESULT 6
US-09-821-821-1

```
; Sequence 1, Application US/09821821
; Patent No. US20020064823A1
; GENERAL INFORMATION:
; APPLICANT: Welcher, Andrew A.
; APPLICANT: Calzone, Frank J.
; TITLE OF INVENTION: CD20/19e-Receptor Like Molecules and Uses Thereof
; FILE REFERENCE: 01017/36938A
; CURRENT APPLICATION NUMBER: US/09/821,821
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: US 09/723,258
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: US 60/193,728
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentl Ver. 2.0
; SEQ ID NO 1
; LENGTH: 760
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (98)..(697)
US-09-821-821-1
```

```
Query Match          99.2%; Score 598.2; DB 10; Length 760;
Best Local Similarity 99.5%; Pred. No. 5,6e-142;
Matches 600; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 ATGATTTCAAGCAGCCGACAGAGTCGGGTCTGTATTTCTCCAGAAATCACTGCT 60
   |||||||
Db 98 ATGATTTCAAGCAGCCGACAGAGTCGGGTCTGTATTTCTCCAGAAATCACTGCT 157
   |||||||
QY 61 TCAGAAATAGTTCACAGAACTTTCAGCAGACCTTTTCAACCAAAAGCCCTTGCAA 120
   |||||||
Db 158 TCAGAAATAGTTCACAGAACTTTCAGCAGACCTTTTCAACCAAAAGCCCTTGCAA 217
   |||||||
QY 121 AATTTTTCGTAGAAATGAAATCTTAGGACTATCCAGATCTCTGTTGGAATTATG 180
   |||||||
Db 218 AATTTTTCGTAGAAATGAAATCTTAGGACTATCCAGATCTCTGTTGGAATTATG 277
   |||||||
QY 181 ACCTTTCTTTGGAGATATCTCTTTCACCTGTAAACCATATCCAGAGTTTCC 240
   |||||||
Db 278 ACCTTTCTTTGGAGATATCTCTTTCACCTGTAAACCATATCCAGAGTTTCC 337
   |||||||
QY 241 TTTTATTTCTTTCAGGATATCCATCTGGGGCTCTGTTTGTCTAATTAATTCGAGGC 300
   |||||||
Db 338 TTTTATTTCTTTCAGGATATCCATCTGGGGCTCTGTTTGTCTAATTAATTCGAGGC 397
   |||||||
QY 301 TTCCTAATTCAGTGAAGAAAGAAACACAGAAACCTGTATATATTTGAGCCGAATATG 360
   |||||||
Db 398 TTCCTAATTCAGTGAAGAAAGAAACACAGAAACCTGTATATATTTGAGCCGAATATG 457
   |||||||
QY 361 AATCTTCTTAGTCCCTGAGAGCAATAGCTGGAATCATCTCTCCACATTTGGTTTATC 420
   |||||||
Db 458 AATTTTCTTAGTCCCTGAGAGCAATAGCTGGAATCATCTCTCCACATTTGGTTTATC 517
   |||||||
QY 421 CTAGATCAAAACTACTATTGTGTATTTCTACCAAAATATGTCAGTGAAGCTGTACT 480
   |||||||
Db 518 CTAGATCAAAACTACTATTGTGTATTTCTACCAAAATATGTCAGTGAAGCTGTACT 577
   |||||||
QY 481 GTCCTGTTCTTGGGAATTTGATTCATATGATGACTTTCAGATTAATTAATTCAT 540
   |||||||
Db 578 GTCCTGTTCTTGGGAATTTGATTCATATGATGACTTTCAGATTAATTAATTCAT 637
   |||||||
QY 541 TCTTGCCTTCTTCAATTTTGGGGTCCACTCAGAGATTTGATTTGAACAATGTTGT 600
   |||||||
Db 638 TCTTGCCTTCTTCAATTTTGGGGTCCACTCAGAGATTTGATTTGAACAATGTTGT 697
   |||||||
QY 601 TGA 603
   |||
Db 698 TGA 700
```

```
RESULT 7
US-09-949-842-10
; Sequence 10, Application US/09949842
; Patent No. US20020164692A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: PT047P1
; FILE REFERENCE: Immune System-related Polynucleotides, Polypeptides, and Antibodies
; CURRENT APPLICATION NUMBER: US/09/949,842
; PRIORITY FILING DATE: 2001-09-02
; PRIOR APPLICATION NUMBER: PCT/US01/07260
; PRIOR FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: 60/224,367
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/187,873
; PRIOR FILING DATE: 2000-03-08
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 689
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-949-842-10

Query Match          97.5%; Score 587.8; DB 9; Length 689;
Best Local Similarity 99.5%; Pred. No. 2.3e-139;
Matches 600; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 ATGGAATTCAGACACCGACACAGTCGCGTGTTCGTGATTTCTCCAGAAATCAGTCT 60
    |||||||
DB 29 ATGATTTCAAGCACCACGACACAGTCGCGTGTTCGTGATTTCTCCAGAAATCAGTCT 88
    |||||||
QY 61 TCAGAAATATGATGTCACAGAACTTTCAGCCAGACCTTTTCACTCAAAAGCCCTTGCAA 120
    |||||||
DB 89 TCAGAAATATGATGTCACAGAACTTTCAGCCAGACCTTTTCACTCAAAAGCCCTTGCAA 148
    |||||||
QY 121 AAATATATTTGCTAGAAAAATGAATCTTAGGGACTATCCAGATCCCTTTGGAAATATG 180
    |||||||
DB 149 AAATATATTTGCTAGAAAAATGAATCTTAGGGACTATCCAGATCCCTTTGGAAATATG 208
    |||||||
QY 181 ACCTTTCTTTTGGAGTTATCTCTTTTCACTTGTAAACCAATATCCAAAGSTTTCC 240
    |||||||
DB 209 ACCTTTCTTTTGGAGTTATCTCTTTTCACTTGTAAACCAATATCCAAAGSTTTCC 268
    |||||||
QY 241 TTTATATTTCTTTCAGATATCATCTCGGGCTCTGTTTGTTCATTAAATTCAGAGCC 300
    |||||||
DB 269 TTTATATTTCTTTCAGATATCATCTCGGGCTCTGTTTGTTCATTAAATTCAGAGCC 328
    |||||||
QY 301 TTCTTATTTGACGAGAAAGAAAAACACAGAACTCTGATTAATTTGAGCCGAATATG 360
    |||||||
DB 329 TTCTTAAATTCAGAGTGAAGAAAAACACAGAACTCTGATTAATTTGAGCCGAATATG 388
    |||||||
QY 361 AATCTCTTAGTGCCTTGAGAGCAATAGCTGGAATCATTTCTCCATCAATTTGGTTTATC 420
    |||||||
DB 389 AATTTTCTTAGTGCCTTGAGAGCAATAGCTGGAATCATTTCTCCATCAATTTGGTTTATC 448
    |||||||
QY 421 CTAGATCAAAAATCTACATTTGTGTATTTCTCACCAAAATAGTCAAGCTGTAAGCTTACT 480
    |||||||
DB 449 CTAGATCAAAAATCTACATTTGTGTATTTCTCACCAAAATAGTCAAGCTGTAAGCTTACT 508
    |||||||
QY 481 GTCTGTGTTCTTGGGAATTTGATTTACATTTGATGACTTTACAGCATTAATTTGATTTACT 540
    |||||||
DB 509 GTCTGTGTTCTTGGGAATTTGATTTACATTTGATGACTTTACAGCATTAATTTGATTTACT 568
    |||||||
QY 541 TCTCTGCTTTCTCAATTTGGGGTGCCACTCAGAGGATTTGATTTGAAACATTTGTTGT 600
    |||||||
DB 569 TCTCTGCTTTCTCAATTTGGGGTGCCACTCAGAGGATTTGATTTGAAACATTTGTTGT 627
    |||||||
QY 601 TGA 603
    |||
DB 628 TGA 630
```

```
RESULT 8
US-09-735-712-7
; Sequence 7, Application US/09735712
; Patent No. US20020045743A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; TITLE OF INVENTION: No. US20020045743A1el Human Membrane Proteins and
; FILE REFERENCE: Lex-0109-USA
; CURRENT APPLICATION NUMBER: US/09/735,712
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 60/171,567
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 450
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-735-712-7

Query Match          57.6%; Score 347.4; DB 10; Length 450;
Best Local Similarity 97.0%; Pred. No. 1.1e-78;
Matches 354; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 ATGGAATTCAGACACCGACACAGTCGCGTGTTCGTGATTTCTCCAGAAATCAGTCT 60
    |||||||
DB 1 ATGGAATTCAGACACCGACACAGTCGCGTGTTCGTGATTTCTCCAGAAATCAGTCT 60
    |||||||
QY 61 TCAGAAATATGATGTCACAGAACTTTCAGCCAGACCTTTTCACTCAAAAGCCCTTGCAA 120
    |||||||
DB 61 TCAGAAATATGATGTCACAGAACTTTCAGCCAGACCTTTTCACTCAAAAGCCCTTGCAA 120
    |||||||
QY 121 AAATATATTTGCTAGAAAAATGAATCTTAGGGACTATCCAGATCCCTTTGGAAATATG 180
    |||||||
DB 121 AAATATATTTGCTAGAAAAATGAATCTTAGGGACTATCCAGATCCCTTTGGAAATATG 180
    |||||||
QY 181 ACCTTTCTTTTGGAGTTATCTCTTTTCACTTGTAAACCAATATCCAAAGSTTTCC 240
    |||||||
DB 181 ACCTTTCTTTTGGAGTTATCTCTTTTCACTTGTAAACCAATATCCAAAGSTTTCC 240
    |||||||
QY 241 TTTATATTTCTTTCAGATATCATCTCGGGCTCTGTTTGTTCATTAAATTCAGAGCC 300
    |||||||
DB 241 TTTATATTTCTTTCAGATATCATCTCGGGCTCTGTTTGTTCATTAAATTCAGAGCC 300
    |||||||
QY 301 TTCTTAAATTCAGAGTGAAGAAAAACACAGAACTCTGATTAATTTGAGCCGAATATG 360
    |||||||
DB 301 TTCTTAAATTCAGAGTGAAGAAAAACACAGAACTCTGATTAATTTGAGCCGAATATG 360
    |||||||
QY 361 AATCT 365
    |||
DB 361 ACTTT 365

RESULT 9
US-09-735-712-5
; Sequence 5, Application US/09735712
; Patent No. US20020045743A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; TITLE OF INVENTION: No. US20020045743A1el Human Membrane Proteins and
; FILE REFERENCE: Lex-0109-USA
; CURRENT APPLICATION NUMBER: US/09/735,712
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 60/171,567
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 417
```

```
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-735-712-5

Query Match
Best Local Similarity 56.7%; Score 341.8; DB 10; Length 417;
Matches 343; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGAATTAAGACCGGACACAGTCCGGTGTTCGTATTTCTCCAGAAATCACTGCT 60
   |||||||
Db 1 ATGGATTAACGACCGGACACAGTCCGGTGTTCGTATTTCTCCAGAAATCACTGCT 60

QY 61 TCAGAAATAGATCCACAGAACTTTCAGCCAGACCTTTTCAACCAAGCCCCCTGGAA 120
   |||||||
Db 61 TCAGAAATAGATCCACAGAACTTTCAGCCAGACCTTTTCAACCAAGCCCCCTGGAA 120

QY 121 AAATTAATTTGCTAGAAAAATGAAATCTTAGGACTATCCAGATCCCTTTGGAAATATG 180
   |||||||
Db 121 AAATTAATTTGCTAGAAAAATGAAATCTTAGGACTATCCAGATCCCTTTGGAAATATG 180

QY 181 ACCTTTCTTTTGAGATATCTTCTCTTTCACTTTGTTAAACCATATCCAAAGTTTCC 240
   |||||||
Db 181 ACCTTTCTTTTGAGATATCTTCTCTTTCACTTTGTTAAACCATATCCAAAGTTTCC 240

QY 241 TTTATATTTCTTTCAGATATCCATCTGGGGCTGTGTTTGTTCATTAATTCGGAGCC 300
   |||||||
Db 241 TTTATATTTCTTTCAGATATCCATCTGGGGCTGTGTTTGTTCATTAATTCGGAGCC 300

QY 301 TTCTTAATTCGAGTGAAGAAAAACACAGAAACTCTGATATA 345
   |||||||
Db 301 TTCTTAATTCGAGTGAAGAAAAACACAGAAACTCTGATATA 345

RESULT 10
US-09-964-824A-207/C
; Sequence 207, Application US/09964824A
; Patent No. US20020102531A1
; GENERAL INFORMATION:
; APPLICANT: Horigan, Stephen
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 689290-73
; CURRENT APPLICATION NUMBER: US/09/964,824A
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/60/236,033
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,032
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,028
; NUMBER OF SEQ ID NOS: 583
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 207
; LENGTH: 382
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-964-824A-207

Query Match
Best Local Similarity 56.4%; Score 340.2; DB 10; Length 382;
Matches 342; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 259 TATTCATTTGGGGCTCTGTTTGTTCATTAATTCGGAGCCCTCCATAATGGAGTAAA 318
   |||||||
Db 382 TATTCATTTGGGGCTCTGTTTGTTCATTAATTCGGAGCCCTCCATAATGGAGTAAA 323

QY 319 AGAAAAACACAGAACTCTGATATAATTTAGCCGATATGAATCTTCTTAGGCCCTG 378
   |||||||
Db 322 AGAAAAACACAGAACTCTGATATAATTTAGCCGATATGAATCTTCTTAGGCCCTG 383

QY 379 AGAGCAATAGCTGGAATCATTTCTCCACATTTGGTTTCAATCCATAGATCAAACTACATT 438
   |||||||
Db 262 GGAGCAATAGCTGGAATCATTTCTCCACATTTGGTTTCAATCCATAGATCAAACTACATT 203
```

```
QY 439 TGTGTTATTTCTACACAAATAGTCAGTGAAGCTGTACTGTCTGTTCGGGAATT 498
   |||||||
Db 202 TGTGTTATTTCTACACAAATAGTCAGTGAAGCTGTACTGTCTGTTCGGGAATT 143

QY 499 TTGATTAATGATGATCTTTCAGCAATTAATGATTAATTCATTTCTGCTTTCATTT 558
   |||||||
Db 142 TTGATTAATGATGATCTTTCAGCAATTAATGATTAATTCATTTCTGCTTTCATTT 83

QY 559 TTGGGCTCCACTCAGAGAAATGATGTTGTGAACAAGTTGTGA 603
   |||||||
Db 82 TTGGGCTCCACTCAGAGAAATGATGTTGTGAACAAGTTGTGA 38

RESULT 11
US-09-735-712-3
; Sequence 3, Application US/09735712
; Patent No. US20020045743A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20020045743A1el Human Membrane Proteins and
; FILE REFERENCE: LEX-0109-USA
; CURRENT APPLICATION NUMBER: US/09/735,712
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 60/171,567
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-735-712-3

Query Match
Best Local Similarity 40.8%; Score 246; DB 10; Length 246;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 358 ATGAATCTTCTTAGTGGCCCTGAGAGCAATAGCTGATCATTTCCCTACATTTGGTTTC 417
   |||||||
Db 1 ATGAATCTTCTTAGTGGCCCTGAGAGCAATAGCTGATCATTTCCCTACATTTGGTTTC 60

QY 418 ATCTGATCAAAACTCATTTTGTGTTATTTCTCACCAAAATAGTCAGTGAAGCTGTT 477
   |||||||
Db 61 ATCTGATCAAAACTCATTTTGTGTTATTTCTCACCAAAATAGTCAGTGAAGCTGTT 120

QY 478 ACGTCTCTGTTCTTGGAAATTTGATTAATGATGATGACTTTCAGCAATTAATATTC 537
   |||||||
Db 121 ACGTCTCTGTTCTTGGAAATTTGATTAATGATGATGACTTTCAGCAATTAATATTC 180

QY 538 ATTCTCTGCTTTCATTAATTTTGGGGTCCACTCAGAGAGATGATGTAAGCAATGT 597
   |||||||
Db 181 ATTCTCTGCTTTCATTAATTTTGGGGTCCACTCAGAGAGATGATGTAAGCAATGT 240

QY 598 TGTGGA 603
   |||||||
Db 241 TGTGGA 246

RESULT 12
US-09-731-872-21
; Sequence 21, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bouquelerec, Lydie
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78 US3 REG
; CURRENT APPLICATION NUMBER: US/09/731,872
; CURRENT FILING DATE: 2000-12-07
```

```
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 21
; LENGTH: 468
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 76..276
; NAME/KEY: sig_peptide
; LOCATION: 76..135
; OTHER INFORMATION: Von Heljne matrix
; OTHER INFORMATION: score 5.21332530399231
; OTHER INFORMATION: seq SPVFLVFPPEITA/SE
US-09-731-872-21
```

```
Query Match          26.3%; Score 158.4; DB 10; Length 468:
Best Local Similarity 75.0%; Pred. No. 6.1e-31;
Matches 198; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
```

```
OY 1 ATGATTTCAAGCACCGCACAGACAGTCCGGTCTTGATATTTCTCCAGAAATCATGCT 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 76 ATGATTTCAAGCACCGCACAGACAGTCCGGTCTTGATATTTCTCCAGAAATCATGCT 135
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 61 TCAGAAATATGATTCACAGAACTTTTCAGCCAGACCTTTTCAACTCAAGACCCCTTGCAA 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 136 TCAGAAATATGATTCACAGAACTTTTCAGCCAGACCTTTTCAACTCAAGACCCCTTGCAA 195
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 121 AATTTTTCCTTGAGAAATGAAATCTTAAAGGACATATCCAGATCCTGTTGGAATTAG 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 196 AATTTTTCCTTGAGAAATGAAATCTTAAAGGACATATCCAGATCCTGTTGTTGT 255
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 181 ACCTTTTCCTTGAGATTAATCTTTCACCTTGTAAACCATATCCAAAGCTTCC 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 256 TCATTAATTCGTGAGCCTTCCTTAATTCGATGGAAGGAAAGGAAAGGAAAGGAA 315
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 241 TTTTATTTCTTTCAGAGATATCCA 264
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 316 TTTTGATTAATTCGATGACTTCA 339
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 13

```
US-09-822-846-227
; Sequence 227, Application US/09822846
; Publication No. US20030027139A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavallee, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Agostino, Michael J.
; APPLICANT: Steininger II, Robert J.
; APPLICANT: Bowman, Michael R.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtel, Kim
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukota, Kamalakara
; APPLICANT: Graham, James R.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6400
; CURRENT APPLICATION NUMBER: US/09/822,846
; CURRENT FILING DATE: 2001-03-29
```

```
; PRIOR APPLICATION NUMBER: 60/195,605
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 629
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 227
; LENGTH: 2438
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-822-846-227
```

```
Query Match          14.3%; Score 86.4; DB 9; Length 2438:
Best Local Similarity 51.1%; Pred. No. 1.6e-12;
Matches 232; Conservative 0; Mismatches 216; Indels 6; Gaps 1;
```

```
OY 119 AAAATTTATTTGCTAGAAAATGAAATCTTAGGACATATCCAGATCCGTTTGAATTA 178
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1610 AGAAGTTCTTGAAGGAGAACCCCAAGTCTTGCGGTGTGTCAGATTCAGCCCTGA 1669
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 179 TGACCTTTCTTCTTGAGATTAATCTTCCCTTTCACCTGTTAAACCATATCCAGATTTG 238
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1670 TGAAGCTTAGCATGGGAAATGATATGATGATGATGATGATGATGATGATGATGATG 1729
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 239 CCTTTATTTCTTTCAGGATTAATCTTTCAGGACATCTTTCAGGACATCTTTCAGGAC 298
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1730 CTAATTCCTGCTATTCAGGATTAATCTTTCAGGACATCTTTCAGGACATCTTTCAGGAC 1789
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 299 CCTTCTTAATTCAGTGAAGAAAGAAACACACAAACTCTGATTAATTTAGCCGAATTA 358
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1790 CCTTGTCAATTTGAGAGAGAAATTAAGAAACAAAGCCGCTGCGAGTATGCTTGAGAA 1849
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 359 TGAATCTTCTTTCAGGACATTAATCTTTCAGGACATCTTTCAGGACATCTTTCAGGAC 418
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1850 TGAATCTTTCAGGACATTAATCTTTCAGGACATCTTTCAGGACATCTTTCAGGAC 1909
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 419 TCCATGAT-----CAAACTACATTTGTTATTTCTCACCAAAATAGTACAGTAAAG 472
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1910 CGTTTATTAATTCATTCACATCACCCTTACGTAATGATGCAATCAATTAATTTGTCATG 1969
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 473 CTGTTCATGCTGCTGCTTCTTGGAATTTTGTATTCATTTGATGATTCAGATTAATTAAT 532
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1970 GGAATATGTCATCTTAATGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2029
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 533 TATTCATTTCTTCTGCTTCTCAATTTTGGGGTG 566
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2030 TCTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2063
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 14

```
US-09-765-205-21
; Sequence 21, Application US/09765205
; Patent No. US20020034800A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Cao, Li
; TITLE OF INVENTION: BONE MARROW SECRETED PROTEINS AND POLYNUCLEOTIDES
; FILE REFERENCE: 1458,004/200130,449
; CURRENT APPLICATION NUMBER: US/09/765,205
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US/09/212,440
; PRIOR FILING DATE: 1998-12-16
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 1563
; TYPE: DNA
; ORGANISM: human
US-09-765-205-21
```

```
Query Match          10.4%; Score 62.6; DB 10; Length 1563:
Best Local Similarity 55.2%; Pred. No. 1.5e-06;
Matches 122; Conservative 0; Mismatches 99; Indels 0; Gaps 0;
OY 119 AAAATTTATTTGCTAGAAAATGAAATCTTAGGACATATCCAGATCCTGTTGGAATTA 178
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Db 364 AGAGTCTTGAAGGAGAGACCAAGTCTGGGGTGTGAGATTCGACTGCCCTGA 423
QY 179 TGACCTTTCTTTGGAGTATCTCTTTCACCTTTTAACCATATCCAGGTTTC 238
Db 424 TGAGCCTTGCATGGGAATACCAATGATGTGTATGCACTATATCTATGGAAGTACC 483
QY 239 CCTTATATTTCTTTCAGATATCCATTCGTGGCTCTGTTTGTTCATTAATTCGAG 298
Db 484 CTATTTCCCTGTATATCCGGGTACACAAATTTGGGGTCACTAATGTTTATTTTCAGAT 543
QY 299 CCTTCCTATTCGAGTGAAGAAAACACAGAAACCTGTG 339
Db 544 CCTGTCAATTCGACGAGGAATTAGAATCAAAAGGCTG 584

RESULT 15

US-09-796-692-6874
Sequence 6874, Application US/09796692
Publication No. US20020198362A1

GENERAL INFORMATION:

APPLICANT: Gaiger, Alexander
APPLICANT: Aligste, Paul A.
APPLICANT: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
FILE REFERENCE: 2077.001200
CURRENT APPLICATION NUMBER: US/09/796,692
PRIOR FILING DATE: 2001-03-01
PRIOR APPLICATION NUMBER: 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/222,903
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: 60/223,416
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: 60/223,378
PRIOR FILING DATE: 2000-08-07
NUMBER OF SEQ ID NOS: 9597
SOFTWARE: fastseq for Windows Version 3.0
SEQ ID NO 6874
LENGTH: 470
TYPE: DNA
ORGANISM: Homo sapiens
US-09-796-692-6874

Query Match 9.1%; Score 55; DB 9; Length 470;

Best Local Similarity 49.9%; Pred. No. 8.1e-05;

Matches 173; Conservative 0; Mismatches 165; Indels 9; Gaps 1;

QY 55 ACTGCTTGAATATGATCCAGACTTTGAGCCAGACCTTTCACTCAAGCCCC 114
Db 91 ATTGCTATGCAATCTGTCACAAACACTCTCAGAGAGATGTCTCACTGGTGGCCCC 150
QY 115 TTGCAAAATTTTGTGTAAGAAATGAAATCTAGGAGCTATGCCATCTCTTTGA 174
Db 151 ACGCAAGCTTCTTATGAGGGAATCTAAGACTTTGGGGGCTGTCCAGATTAATGAGG 210
QY 175 ATTATGACCTTTCTTTGAGATATCTCTTTACCTTGTAAACATATCCAAAG 234

Db 211 CTCTTCACATTCGCCCTGGGGGCTCTTGATGATCCAGCAGGATCTATGCACCCATC 270
QY 235 TTTCCTTTATTTTCTTTTCAGGATATCCATTCGTGGGCTCTGTTTGTTCATTAATCT 294
Db 271 TGTGTGACTGTGT-----GTACCTCTCTGTGGGAGCATTAATATTTTC 321
QY 295 GGAGCCTTCTTAATTCAGATGAAGAAAACACAGAAACTGTATATATTTGAGCCGA 354
Db 322 GGATCACTCTCGCAGCAACGAGAAAACCTCAGGAAGTGTGTGTCAAAGAAAATG 381
QY 355 ATAATGAATCTTCTTAGTCCCTGAGAGCAATAGCTGGAATATCT 401
Db 382 AATAATGAATTCATTAGCCCTCTTGTGTCATTCGTGAATGATTC 428

Search completed: February 22, 2003, 12:17:45
Job time : 60 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 22, 2003, 12:20:16 ; Search time 2199 Seconds

(without alignments)
4441.054 Million cell updates/sec

Title: US-09-735-712-1

Perfect score: 603
Sequence: 1 atgattcaagcagccagcaca.....attgtgaacatgttgtga 603

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estinu:*
5: em_estov:*
6: em_estipl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	598.8	99.3	739	14 BQ429208	BQ429208 AGENCOURT
2	595.6	98.8	713	13 B1829760	B1829760 603079831
3	466.2	77.3	508	9 A1149899	A1149899 q143n06.x
4	450.6	74.7	516	9 AA436088	AA436088 zu03a08.x
5	347	57.5	387	9 AA416972	AA416972 zt94n05.s
6	340.2	56.4	382	9 AA435988	AA435988 zu03a08.s

C	7	291.8	48.4	394	9	AA758635
C	8	283.8	47.1	484	10	BE107659
C	9	267.4	44.3	410	9	AA411806
C	10	245.4	40.7	389	9	AA781801
C	11	242.8	40.3	415	9	AA470059
C	12	207	34.3	482	9	A1002083
C	13	141	23.4	332	12	BF319786
C	14	131.8	21.9	601	10	BE638317
C	15	126.2	20.9	224	10	BE638325
C	16	115.6	19.2	538	17	AQ108532
C	17	113	18.7	281	9	AA707529
C	18	96.6	16.0	568	13	BM253528
C	19	86.4	14.3	790	13	BT771845
C	20	86.4	14.3	1071	9	AL544561
C	21	85	14.1	900	13	BT759693
C	22	84.8	14.1	581	14	BM713748
C	23	84	13.9	793	12	BG484817
C	24	83	13.8	218	10	BE638321
C	25	82.8	13.7	799	13	BT908709
C	26	82.2	13.6	704	12	BG571626
C	27	80.8	13.4	517	12	BG664654
C	28	80.4	13.3	508	9	AA344138
C	29	80.4	13.3	562	9	AA418443
C	30	76.2	12.6	857	12	BG720182
C	31	74.4	12.3	1090	14	BM919064
C	32	73	12.1	713	13	BT562282
C	33	70.6	11.7	421	17	BG6842
C	34	65.4	10.8	736	14	BQ43535
C	35	65.4	10.8	921	13	BT561611
C	36	65.4	10.8	998	13	BM554322
C	37	64.4	10.7	775	13	BT560610
C	38	63.6	10.5	689	12	BG719295
C	39	63.4	10.5	856	13	BT560592
C	40	63.4	10.5	899	12	BT560592
C	41	61.2	10.1	887	9	AL531049
C	42	59.4	9.9	715	12	BG538851
C	43	59.4	9.9	1157	13	BM552706
C	44	55.8	9.3	656	14	BM781186
C	45	55.6	9.2	579	10	BE513276

ALIGNMENTS

RESULT 1
BQ429208
LOCUS BQ429208 739 bp mRNA linear EST 24-MAY-2002
DEFINITION AGENCOURT_7836975 NIH_MGC_82 Homo sapiens cDNA clone IMAGE:6102065
5', mRNA sequence.
ACCESSION BQ429208
VERSION BQ429208.1 GI:21168284
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 739)
NIH-MGC <http://mgc.ncl.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strusberg, Ph.D.
Email: cgephs-remail.nih.gov
Tissue Procurement: CLONTECH
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNU)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNU at:
<http://image.llnl.gov>
Plate: L10C2336 row: b column: 18
High quality sequence stop: 544.
Location/Qualifiers
1. 739

FEATURES
source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:610265"
/clone_lib="NIH_MGC_82"
/lab_host="DH10B (TI phage-resistant)"
/notes="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1:
5' (ggccgcctcgcc); Site_2: 5' (ggccatctagcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CAGGCGCATATGCGC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGGAGCGGCGCCACATG-df(30)-BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size
1.35 kb (range 0.9-4.0 kb). 14/15 colonies contained
inserts by PCR. This library was enriched for full-length
clones and was constructed by Clontech Laboratories (Palo
Alto, CA)."
```

BASE COUNT 204 a 148 c 126 g 239 t 22 others

ORIGIN

Query Match 99.3%; Score 598.8; DB 14; Length 739;
Best Local Similarity 99.5%; Pred. No. 7.8e-134;
Matches 600; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

OY 1 ATGGATTCAAGCACCACACAGTCGGGTCTTCTGTATTTCTCCACAAATCACTGCT 60
DB 74 ATGGATTCAAGCACCACACAGTCGGGTCTTCTGTATTTCTCCACAAATCACTGCT 133
OY 61 TCAGAAATATGAGTCACAGAACTTTCAGCAGACCTTTCAACTCAAGCCCTTGCAA 120
DB 134 TCAGAAATATGAGTCACAGAACTTTCAGCAGACCTTTCAACTCAAGCCCTTGCAA 193
OY 121 AAATATTTGCTAGAAAAATGAAATCTTAGGACTATCCAGATCCTGTTGGAAATTAG 180
DB 194 AAATATTTGCTAGAAAAATGAAATCTTAGGACTATCCAGATCCTGTTGGAAATTAG 253
OY 181 ACCCTTTCTTTGGAGTATCTTCTCTTCACTCTTTAAACCAATCCAGAGTTTCCC 240
DB 254 ACCCTTTCTTTGGAGTATCTTCTCTTCACTCTTTAAACCAATCCAGAGTTTCCC 313
OY 241 TTTATATTTCTTTCAGAGATATCCATCTGGGCTCTGTTTGTATTTCTGAGGCG 300
DB 314 TTTATATTTCTTTCAGAGATATCCATCTGGGCTCTGTTTGTATTTCTGAGGCG 373
OY 301 TTCCTAATTCAGTAAAAAGAAACACAGAACTCTGATTAATTTGACCGAATTATG 360
DB 374 TTCCTAATTCAGTAAAAAGAAACACAGAACTCTGATTAATTTGACCGAATTATG 433
OY 361 AATCTCTTAGTCCCTGAGACAATAGCTGGAAATCATCTCCACATTTGGTTTCATC 420
DB 434 AATCTCTTAGTCCCTGAGACAATAGCTGGAAATCATCTCCACATTTGGTTTCATC 493
OY 421 CTAGATCAAAACTACATTTGTGTTATTTCTCACCAAAATAGTCAGTAAAGCTGTACT 480
DB 494 CTAGATCAAAACTACATTTGTGTTATTTCTCACCAAAATAGTCAGTAAAGCTGTACT 553
OY 481 GTCCGTCTCTTGGGAATTTTGATTTGATTTGATGACTTTAGACATTAATTAATTCATT 540
DB 554 GTCCGTCTCTTGGGAATTTTGATTTGATTTGATGACTTTAGACATTAATTAATTCATT 613
OY 541 TCTTGCCCTTCTCAATTTTGGGTGCGCATCAGAGAGATTTGATTTGTAACAATTTGT 600
DB 614 TCTTGCCCTTCTCAATTTTGGGTGCGCATCAGAGAGATTTGATTTGTAACAATTTGT 673
OY 601 TGA 603
DB 674 TGA 676
```

RESULT 2
LOCUS B1829760 713 bp mRNA linear EST 04-OCT-2001
DEFINITION 603079831P1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5171526 5',
ACCESSION B1829760 mRNA sequence.

```

VERSION B1829760.1 GI:15941310
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Eumalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgc.ncl.nih.gov/.
1 (bases 1 to 713)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNI)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMNI at:
http://image.llnl.gov
Plate: LLM11426 row: n column: 07
High quality sequence stop: 713.
Location/Qualifiers
1..713
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5171526"
/clone_lib="NIH_MGC_119"
/tissue_type="medulla"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb. Insert size range
0.9-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invitrogen). Research Genetics tracking code 013. Note:
this is a NIH-MGC Library."
```

BASE COUNT 202 a 154 c 123 g 233 t 1 others

ORIGIN

Query Match 98.8%; Score 595.6; DB 13; Length 713;
Best Local Similarity 99.2%; Pred. No. 4.6e-133;
Matches 598; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

OY 1 ATGGATTCAAGCACCACACAGTCGGGTCTTCTGTATTTCTCCAGAAATCACTGCT 60
DB 87 ATGGATTCAAGCACCACACAGTCGGGTCTTCTGTATTTCTCCAGAAATCACTGCT 146
OY 61 TCAGAAATATGAGTCACAGAACTTTCAGCAGACCTTTCAACTCAAGCCCTTGCAA 120
DB 147 TCAGAAATATGAGTCACAGAACTTTCAGCAGACCTTTCAACTCAAGCCCTTGCAA 206
OY 121 AAATATTTGCTAGAAAAATGAAATCTTAGGACTATCCAGATCCTGTTGGAAATTAG 180
DB 207 AAATATTTGCTAGAAAAATGAAATCTTAGGACTATCCAGATCCTGTTGGAAATTAG 266
OY 181 ACCCTTTCTTTGGAGTATCTTCTCTTCACTCTTTAAACCAATCCAGAGTTTCCC 240
DB 267 ACCCTTTCTTTGGAGTATCTTCTCTTCACTCTTTAAACCAATCCAGAGTTTCCC 326
OY 241 TTTATATTTCTTTCAGAGATATCCATCTGGGCTCTGTTTGTATTTCTGAGGCGC 300
DB 327 TTTATATTTCTTTCAGAGATATCCATCTGGGCTCTGTTTGTATTTCTGAGGCGC 386
OY 301 TTCCTAATTCAGTAAAAAGAAACACAGAACTCTGATTAATTTGACCGAATTATG 360
DB 387 TTCCTAATTCAGTAAAAAGAAACACAGAACTCTGATTAATTTGACCGAATTATG 446
OY 361 AATCTCTTAGTCCCTGAGACAATAGCTGGAAATCATCTCCACATTTGGTTTCATC 420
DB 447 AATCTCTTAGTCCCTGAGACAATAGCTGGAAATCATCTCCACATTTGGTTTCATC 506
```

```
QY 421 CTAGATCAAACTACATTGTGTGTTATCTGCACAAATAGTCAGTAAAGCTGTACT 480
|||||
Db 507 CTGATCAAACTACATTGTGTGTTATCTGCACAAATAGTCAGTAAAGCTGTACT 566
QY 481 GTCCTTCTTGGGAATTTTGATTGATGACCTTTAGCATTTTGAATTTATTCATT 540
|||||
Db 567 GTCCTTCTTGGGAATTTTGATTGATGACCTTTAGCATTTTGAATTTATTCATT 626
QY 541 TCTCTCCTTTTCATATTTGGGGTGCACATCAGAGATTTGATGTGAACAATGTCT 600
|||||
Db 627 TCTCTCCTTTTCATATTTGGGGTGCACATCAGAGATTTGATGTGAACAATGTCT 686
QY 601 TGA 603
|||||
Db 687 TGA 689

RESULT 3
A1149899/c 508 bp mRNA linear EST 10-NOV-1998
LOCUS g14h06.x1 Soares testis_NHT Homo sapiens cDNA clone IMAGE:1752827
DEFINITION 3', similar to SW:CD20_HUMAN P11836 B-LYMPHOCTE ANTIGEN CD20 ;,
mRNA sequence.
ACCESSION A1149899
VERSION A1149899
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CCAP),
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
, Ph.D.
' cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCAP clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNL at:
www.bio.lnlnl.gov/bbrp/image/image.html
Insert Length: 742 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 455.
FEATURES
source
1..508
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="1752827"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: pRTT3D-Pac (Pharmacia) with a modified
polylinker. Site.1: Not I; Site.2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCCGCAATTTTCTTTTCTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pRTT3 vector. Library
went through one round of normalized to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 186 a 91 c 89 g 142 t
ORIGIN

Query Match 77.3%; Score 466.2; DB 9; Length 508;
Best Local Similarity 99.4%; Pred. No. 6.3e-102;
Matches 468; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 133 AGAAAAATGAAATCTTAGGACATTCAGATCTGTTTGAATATATGACCTTTCTTTT 192
|||||
Db 508 AGAAAAATGAAATCTTAGGACATTCAGATCTGTTTGAATATATGACCTTTCTTTT 449
QY 193 GGAGTTATCTCTCTTTTACCTTTGTTAAACCATTCACAAAGTTTCCCTTATATCTT 252
|||||
Db 448 GGAGTTATCTCTCTTTTACCTTTGTTAAACCATTCACAAAGTTTCCCTTATATCTT 389
QY 253 TCAGATATTCATTTCTGGGGCTCTGTTTGTTCATTAATTTGGAGCTTCCATTTGA 312
|||||
Db 388 TCAGATATTCATTTCTGGGGCTCTGTTTGTTCATTAATTTGGAGCTTCCATTTGA 329
QY 313 GTGAAAAAGAAAAACACAGAAACTGATTAATTTAGCGAATTAATGATCTTCTAGT 372
|||||
Db 328 GTGAAAAAGAAAAACACAGAAACTGATTAATTTAGCGAATTAATGATCTTCTAGT 269
QY 373 GCCCTGAGCAATAGCTGGAATCTTCCTTCACATTTGGTTTCATCTAGATCAAAAC 432
|||||
Db 268 GCCCTGAGCAATAGCTGGAATCTTCCTTCACATTTGGTTTCATCTAGATCAAAAC 209
QY 433 TCATTTGTGTTATCTTCACCAAAATAGTCAGTGAAGCTGTCTGCTGCTTG 492
|||||
Db 208 TCATTTGTGTTATCTTCACCAAAATAGTCAGTGAAGCTGTCTGCTGCTTG 149
QY 493 GGAATTTGATTACATTTGATGACTTTACACATTAATTAATTCATTTCTGCTTTC 552
|||||
Db 148 GGAATTTGATTACATTTGATGACTTTACACATTAATTAATTCATTTCTGCTTTC 89
QY 553 TCATTTTGGGGTGCCACATCAGAGATTTGATGTGAACAATGTTGTTGA 603
|||||
Db 88 TCATTTTGGGGTGCCACATCAGAGATTTGATGTGAACAATGTTGTTGA 38

RESULT 4
A436088 516 bp mRNA linear EST 09-NOV-1997
LOCUS zu03a08.r1 Soares testis_NHT Homo sapiens cDNA clone IMAGE:730742
DEFINITION 5', similar to SW:CD20_HUMAN P11836 B-LYMPHOCTE ANTIGEN CD20 ;,
mRNA sequence.
ACCESSION A436088
VERSION A436088.1 GI:2141002
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 516)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Kritzman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
,J., Moore,B., Schellenberg,K., Stepec,M., Tan,F., Theising,B.,
White,X., Wylie,T., Waterston,R. and Wilson,R.
JOURNAL WashU-NCI human EST Project
COMMENT Unpublished (1997)
CONTACT: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through ILNL; contact the
IMAGE Consortium (info@image.lnlnl.gov) for further information.
Insert Length: 733 Std Error: 0.00
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 492.
FEATURES
source
1..516
/organism="Homo sapiens"
/db_xref="GDB:5927437"
/db_xref="taxon:9606"
/clone_image="730742"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: pRTT3D-Pac (Pharmacia) with a modified
```

polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGCGCCGCCCAATTTTCTTTTCTTTT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization to Cot5, and was constructed by Bento Soares and M. Fatima Bonaldo.

BASE COUNT 149 a 117 c 80 g 170 t
ORIGIN

Query Match 74.7%; Score 450.6; DB 9; Length 516;
Best Local Similarity 99.1%; Pred. No. 3.6e-98;
Matches 453; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGATTTAAGCAGCCGACACAGTCCGCTTTCTGTAATTTCTCCAGAAATCACTGCT 60
DB 52 ATGATTTAAGCAGCCGACACAGTCCGCTTTCTGTAATTTCTCCAGAAATCACTGCT 111
QY 61 TCAGAAATAGAGTCCAGAACTTTCAGCAGACCTTTTCAACCAAGCCCTTGCA 120
DB 112 TCAGAAATAGAGTCCAGAACTTTCAGCAGACCTTTTCAACCAAGCCCTTGCA 171
QY 121 AAATTAATTTGCTAGAAAAATGAATCTTAGGACTATCCAGATCCTGTTGGAATTATG 180
DB 172 AAATTAATTTGCTAGAAAAATGAATCTTAGGACTATCCAGATCCTGTTGGAATTATG 231
QY 181 ACCTTTCTTTGGAGTATCTCTTTTCACTTTTAAACCAATTCAGAGTTTCC 240
DB 232 ACCTTTCTTTGGAGTATCTCTTTTCACTTTTAAACCAATTCAGAGTTTCC 291
QY 241 TTTATATTTCTTTCAGAAATCCATTTCTGGGCTCTGTTTGTATTAATTAATCTGAGGC 300
DB 292 TTTATATTTCTTTCAGAAATCCATTTCTGGGCTCTGTTTGTATTAATTAATCTGAGGC 351
QY 301 TTCTTAATTTGCTAGAAAAAGAAACCAAGAACTTGATTAATTTGAGCCGAATATG 360
DB 352 TTCTTAATTTGCTAGAAAAAGAAACCAAGAACTTGATTAATTTGAGCCGAATATG 411
QY 361 AATCTTCTTAGAGCCCTGAGCAATAGCTGGAATCATTCCTCACAATTTGGTTTCATC 420
DB 412 AATTTCTTAGAGCCCTGAGCAATAGCTGGAATCATTCCTCACAATTTGGTTTCATC 471
QY 421 CTAGATCAAAACTACATTTGTGGTTATTCACCAAA 457
DB 472 CTAGATCAAAACTACATTTGTGGTTATTCACCAAA 508

RESULT 5
AA416972 387 bp mRNA linear EST 09-NOV-1997
LOCUS z194h05.s1 Soares testis NHT Homo sapiens cDNA clone IMAGE:730041
DEFINITION 3' similar to SW:CD20_HUMAN P11836 B-LYMPHOCYTE ANTIGEN CD20 ;,
mRNA sequence.

ACCESSION AA416972
VERSION AA416972.1 GI:2077080
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 387)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theisling,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.

TITLE
JOURNAL
COMMENT
Washo-NCI human EST Project
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800

Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 556 Std Error: 0.00
Seq primer: -41m3 fwd, 5' from Amersham.

FEATURES
source location/Qualifiers
1..387

/organism="Homo sapiens"
/db_xref="GDB:5926570"
/db_xref="taxon:9606"
/clone="IMAGE:730041"
/clone_id="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"

/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGCGCCGCCCAATTTTCTTTTCTTTT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization to Cot5, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 138 a 77 c 67 g 105 t
ORIGIN

Query Match 57.5%; Score 347; DB 9; Length 387;
Best Local Similarity 98.6%; Pred. No. 3e-73;
Matches 350; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 249 TCTTACAGAAATCCATTTCTGGGCTCTGTTTGTTCATTAATTCGAGCCTTCTAAT 308
DB 387 TCTTACAGAAATCCATTTCTGGGCTCTGTTTGTTCATTAATTCGAGCCTTCTAAT 328
QY 309 TGCAGTGAAGAAAAACCAAGAACTCTGATTAATTTGAGCCGAATTAATCTTCT 368
DB 327 TGCAGTGAAGAAAAACCAAGAACTCTGATTAATTTGAGCCGAATTAATCTTCT 268
QY 369 TAGTCCCTGAGAGCAATAGCTGGAATCATTCCTCACAATTTGGTTTCATCTGATCA 428
DB 267 TAGTCCCTGAGAGCAATAGCTGGAATCATTCCTCACAATTTGGTTTCATCTGATCA 208
QY 429 AAACATCAATTTGTGTTATTTCTACCAAAATAGTCAGTGAAGCTGTACTGCTGTT 488
DB 207 AAACATCAATTTGTGTTATTTCTACCAAAATAGTCAGTGAAGCTGTACTGCTGTT 148
QY 489 CTGGGAATTTGATTACATTTGATGACATTTTCAATTTATTCATTTCTCTGCC 548
DB 147 CTGGGAATTTGATTACATTTGATGACATTTTCAATTTATTCATTTCTCTGCC 88
QY 549 TTTCATATTTTGGGGTCCATCAGAGATGATGATGTAACAATGTTGTA 603
DB 87 TTTCATATTTTGGGGTCCATCAGAGATGATGATGTAACAATGTTGTA 33

RESULT 6
AA435988 382 bp mRNA linear EST 09-NOV-1997
LOCUS z034a08.s1 Soares testis NHT Homo sapiens cDNA clone IMAGE:730742
DEFINITION 3' similar to SW:CD20_HUMAN P11836 B-LYMPHOCYTE ANTIGEN CD20 ;,
mRNA sequence.

ACCESSION AA435988
VERSION AA435988.1 GI:2140902
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 382)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theisling,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.

RESULT 8
 BE107659/c
 LOCUS
 DEFINITION BE107659 484 bp mRNA linear EST 13-JUN-2000
 UI-R-BT1-ame-d-04-0-UI.s1 UI-R-BT1 Rattus norvegicus cDNA clone
 ACCESSION BE107659
 VERSION BE107659.1 GI:8499769
 KEYWORDS EST.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 484)
 Ronaldo,M.F., Lennon,G. and Soares,M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 Genome Res. 6 (9), 791-806 (1996)
 9704447
 JOURNAL
 MEDLINE
 COMMENT
 Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: msoares@blue.weeg.uiowa.edu
 The sequence contained an oligo-dT track that was present in the
 oligonucleotide that was used to prime the synthesis of first
 strand cDNA and therefore this may represent a bonafide poly A
 tail. The sequence tag present in the cDNA between the NotI site
 and the oligo-dT track served to identify it as a clone from the
 normalized testis library cDNA library Preparation: M.B. Soares Lab
 Clone distribution: clones will be available through Research
 Genetics (www.resgen.com)
 Seq primer: M13 forward
 POLYA=yes.
 FEATURES
 source
 Location/Qualifiers
 1..484
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-BT1-ame-d-04-0-UI"
 /clone_lib="UI-R-BT1"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker. Site_1: Not I; Site_2: Eco RI; The library
 UI-R-BT1 is a subtracted library derived from a mixture of
 the following tissues: hippocampus, thalamus, mid-brain,
 medulla, corpus striatum, cerebral cortex and testis. For
 a detailed description of the library from which this
 clone was derived, please visit our web site at
 ratest.eng.uiowa.edu. The subtraction has been previously
 described in (Ronaldo, Lennon and Soares, Genome Research
 6:791-806, 1996)
 TAG_LIB=UI-R-BT1
 TAG_TISSUE=testis
 TAG_SEQ=ACCGAG"
 BASE COUNT 165 a 98 c 96 g 125 t
 ORIGIN
 Query Match 47.1%; Score 283.8; DB 10; Length 484;
 Best Local Similarity 79.4%; Pred. No. 4.7e-58;
 Matches 336; Conservative 0; Mismatches 87; Indels 0; Gaps 0;
 QY 161 AGATCGTTGGATTTGACCTTTCTTTGGATATCTTCTTACCTGTGTA 220
 |||||||
 DB 484 AGATCGTTGGATTCAGACTTCATTTGGAGTGTTCCTTTACCTGTGTA 425
 |||||||
 QY 221 AACCATTCACAGTTTCCCTTATATTTCTTTCAGATTCATTCGGGGCTGT 280
 |||||||
 DB 424 ACCCATACCAAGTTCCTTTATATTTCTCTCAGATATCTTTCGGGGCTGT 365
 |||||||

QY 281 TGTTCATTATTCGTGAGCTTCTCTAATTCAGTGAAGAAAACACAGAACTCTGA 340
 |||||||
 DB 364 TGTTCATTACTCTGGAGCTTCTTGATTTGCTGTGAAAAGAACTACAGACTCTGA 305
 |||||||
 QY 341 TAATATTGAGCCGAATGAATCTTCTTATGTCCTCGAGCAATACCTGGAATCATTC 400
 |||||||
 DB 304 TAAATATGAGCCAGCGAGATTAATTTACTTATGTCCTCGAGCAAGCTGGAATCATTC 245
 |||||||
 QY 401 TCCCTACATTTGGTTTCATCCCTAGATCAAACTACATTTGCTATTCACCAAAATA 460
 |||||||
 DB 244 TCCCTATATGTCCTTCTCTTCTGATGAGAAATTCATCTGCTATTCACGATGTA 185
 |||||||
 QY 461 GTTCAGTGAAGCTGTATCTGCTGCTTCTTCTGGAATTTGATTCATGATCACTTCA 520
 |||||||
 DB 184 TTCAGTGTGCTCTATTCACATCTATTTGTTGGATTTGATTAATGATGATCTTCA 125
 |||||||
 QY 521 GCATTATTCATTTATTCATTTCTCTGCTTCTTCAATTTTGGGGTCCACTAGAGATT 580
 |||||||
 DB 124 GCGTTCAGACTGTTCATTTCCCTTTCTCGATTTTGGGGTCTACTCAGAAAGA 65
 |||||||
 QY 581 GNG 583
 ||||
 DB 64 GTG 62
 RESULT 9
 AA411806/c
 LOCUS
 DEFINITION AA411806 410 bp mRNA linear EST 12-AUG-1997
 zt67a03.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:727372
 3' similar to SW:CD20_HUMAN P11836 B-LYMPHOCYTE ANTIGEN CD20 ;,
 mRNA sequence.
 ACCESSION AA411806
 VERSION AA411806.1 GI:2070377
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 410)
 Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
 Kucaba,T., Lacy,M., Le,N., Lennon,G., Mariz,M., Martin,D., Moore,B.,
 Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wyllie
 ,T., Waterston,R. and Wilson,R.
 WashU-Merck EST Project 1997
 Unpublished (1997)
 TITLE JOURNAL
 COMMENT
 Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 608 Std Error: 0.00
 Seq primer: -41m13 fwd. Ef from Amersham
 High quality sequence stop: 401.
 FEATURES
 source
 Location/Qualifiers
 1..410
 /organism="Homo sapiens"
 /db_xref="GDB:5924284"
 /db_xref="taxon:9606"
 /clone="IMAGE:727372"
 /clone_lib="Soares_testis_NHT"
 /sex="male"
 /lab_host="DH10B"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was prepared from mRNA obtained from Clontech Laboratories
 , Inc. and primed with a Not I - oligo(dT) primer [5'
 TGTTCACATCTGAGTGGAGCGCGCCCAATTTTTTTTTTTT 3']
 Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I

/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCGCCCAATTTTCTTTTCTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 129 a 96 c 67 g 123 t
ORIGIN

Query Match 40.3%; Score 242.8; DB 9; Length 415;
Best Local Similarity 77.3%; Pred. No. 3.4e-48;
Matches 357; Conservative 0; Mismatches 2; Indels 103; Gaps 1;

Y 1 ATGATTCAGACCGCACACAGTCCGGTGTTCGTATTCCTCCAGAAATCACTGCT 60
|||||
Db 57 ATGATTCAGACCGCACACAGTCCGGTGTTCGTATTCCTCCAGAAATCACTGCT 116
|||||
Qy 61 TCAGAAATAGAGTCCAGCAAACTTTCAGCCAGACCTTTTCACACCAAGCCCTTGCAA 120
Db 117 TCAGAAATAGAGTCCAGCAAACTTTCAGCCAGACCTTTTCACACCAAGCCCTTGCAA 176
|||||
Qy 121 AAATATTGTCTAGAAAATGAATCTTAGGAGTATCCAGATCCTGTTTGGAATTATG 180
|||||
Db 177 AAATATTGTCTAGAAAATGAATCTTAGGAGTATCCAGATCCTGTTTGGAATTATG 208
|||||
Qy 181 ACCTTTCTTTGGAGTATCTCTCTTTCACCTGTTTAAACATATCCAAAGTTTCCC 240
Db 209 ----- 208
Qy 241 TTTATATTCTTTCAGAAATTCATTCGGGCTCTGTTTGTTCATTAATTCGGAGCC 300
|||||
Db 209 -----GGAATTCATTCGGGCTCTGTTTGTTCATTAATTCGGAGCC 253
|||||
Qy 301 TTCTTAATTCAGTGAAGAAAAGAAAACACAGAACTGTGAATATTAGCCGGAATATG 360
Db 254 TTCTTAATTCAGTGAAGAAAAGAAAACACAGAACTGTGAATATTAGCCGGAATATG 313
|||||
Qy 361 AATCTTTTGTAGTCCCTGAGAGCAATAGCTGGAATCATCTTCCTCACAATTTGGTTTCATC 420
Db 314 AATCTTTTGTAGTCCCTGAGAGCAATAGCTGGAATCATCTTCCTCACAATTTGGTTTCATC 373
|||||
Qy 421 CTAGATCAAACTACATTTGTGTTATTCCTACCAAAATATG 462
Db 374 CTAGATCAAACTACATTTGTGTTATTCCTACCAAAATATG 415
|||||

RESULT 12
AI002083 482 bp mRNA linear EST 27-AUG-1998
LOCUS ot38b02.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1619019
DEFINITION 3', mRNA sequence.
ACCESSION AI002083
VERSION AI002083.1 GI:3202120
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 482)
NCI-CCAP <http://www.ncbi.nlm.nih.gov/nciccapd>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-rt@mail.nih.gov
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCAP clone distribution information can be
found through the I.M.A.G.E. Consortium/ILMIL at:
www-bio.liml.gov/bhrp/image/image.html
Insert length: 598 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 393.
Location/Qualifiers
1. .482
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1619019"
/clone_id="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"

BASE COUNT 157 a 85 c 99 g 141 t
ORIGIN

Query Match 34.3%; Score 207; DB 9; Length 482;
Best Local Similarity 68.8%; Pred. No. 1.4e-35;
Matches 415; Conservative 0; Mismatches 0; Indels 188; Gaps 2;

Qy 1 ATGATTCAGACCGCACACAGTCCGGTGTTCGTATTCCTCCAGAAATCACTGCT 60
|||||
Db 450 ATGATTCAGACCGCACACAGTCCGGTGTTCGTATTCCTCCAGAAATCACTGCT 391
|||||
Qy 61 TCAGAAATAGAGTCCAGCAAACTTTCAGCCAGACCTTTTCACACCAAGCCCTTGCAA 120
Db 390 TCAGAAATAGAGTCCAGCAAACTTTCAGCCAGACCTTTTCACACCAAGCCCTTGCAA 331
|||||
Qy 121 AAATATTGTCTAGAAAATGAATCTTAGGAGTATCCAGATCCTGTTTGGAATTATG 180
|||||
Db 330 AAATATTGTCTAGAAAATGAATCTTAGGAGTATCCAGATCCTGTTTGGAATTATG 299
|||||
Qy 181 ACCTTTCTTTGGAGTATCTCTCTTTCACCTGTTTAAACATATCCAAAGTTTCCC 240
Db 298 ----- 299
Qy 241 TTTATATTCTTTCAGAAATTCATTCGGGCTCTGTTTGTTCATTAATTCGGAGCC 300
|||||
Db 298 -----GGAATTCATTCGGGCTCTGTTTGTTCATTAATTCGGAGCC 254
|||||
Qy 301 TTCTTAATTCAGTGAAGAAAAGAAAACACAGAACTGTGAATATTAGCCGGAATATG 360
Db 253 TTCTTAATTCAGTGAAGAAAAGAAAACACAGAACTGTGAATATTAGCCGGAATATG 216
|||||
Qy 361 AATCTTTTGTAGTCCCTGAGAGCAATAGCTGGAATCATCTTCCTCACAATTTGGTTTCATC 420
Db 215 ----- 216
Qy 421 CTAGATCAAACTACATTTGTGTTATTCCTACCAAAATATGCTGAAGCTGTACT 480
|||||
Db 215 ---GATCAAACTACATTTGTGTTATTCCTACCAAAATATGCTGAAGCTGTACT 159
|||||
Qy 481 GTCCGTCTCTGGGAATTTGATTACATTTGACTTTCACATTTATGAAATTTACTT 540
Db 158 GTCCGTCTCTGGGAATTTGATTACATTTGACTTTCACATTTATGAAATTTACTT 99
|||||
Qy 541 TCTCGCTTTCTCAATTTTGGGGTGCACACAGAGATTTGATTTGAAACAATGTGT 600
Db 98 TCTCGCTTTCTCAATTTTGGGGTGCACACAGAGATTTGATTTGAAACAATGTGT 39
|||||
Qy 601 TGA 603

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd

OM protein - protein search, using sw model

Run on: February 24, 2003, 12:59:26 ; Search time 38.3132 Seconds
(without alignments)
692.108 Million cell updates/sec

Title: US-09-735-712-2

Perfect score: 1015
Sequence: 1 MDSTAHSPVFLVFPPETA.....ISLPFSLGCHSEDDCEQC 199

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%

1:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT *
2:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT *
3:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT *
4:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT *
5:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT *
6:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT *
7:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT *
8:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT *
9:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT *
10:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT *
11:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT *
12:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT *
13:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT *
14:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT *
15:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT *
16:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT *
17:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT *
18:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT *
19:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT *
20:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT *
21:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT *
22:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT *
23:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query				ID	Description
No.	Score	Match	Length	DB		
1	1015	100	0	199	22	AAU01210
2	1004	98.9	200	22	22	AAE13062
3	1004	98.9	200	22	22	AA689138
4	923	90.9	201	22	22	AAE10917
5	923	90.9	212	22	22	ABB95818
6	923	90.9	212	22	22	AAW95114
7	726.5	71.6	149	22	22	AAU01213
8	673	66.3	158	22	22	ABB12234
9	583	57.4	138	22	22	AAU01212
10	416	41.0	81	22	22	AAU01211

11	294	29.0	67	21	AAG693074	Human secreted pro
12	263.5	26.0	77	22	AAG693074	Human secreted pro
13	206.5	20.3	220	23	ABBS00317	Human polypeptide
14	206.5	20.3	239	29	AAW66745	High affinity immu
15	206.5	20.3	239	21	AAW501174	Human high affinity
16	206.5	20.3	245	21	AAW94973	Human secreted pro
17	205.5	20.2	220	22	AAW83512	Human polypeptide,
18	184.5	18.2	214	17	AAW06503	Httm protein, Hom
19	184.5	18.2	214	19	AAW41056	Httm protein, Hom
20	184.5	18.2	225	21	AAW94449	Human inflammation
21	184.5	18.2	242	21	AAW34495	Human secreted pro
22	184.5	18.2	248	20	AAW15225	Human receptor pro
23	184.5	18.2	248	21	AAW91531	Human secreted pro
24	184.5	18.2	248	22	AAE12072	Dendritic cell (DC
25	184.5	18.2	248	22	AAW07489	Human hHAERBs-iso
26	184.5	18.2	248	23	ABBS00341	Human polypeptide
27	184.5	18.2	250	20	AAW84505	Human breast tumou
28	184.5	18.2	257	21	AAW58419	Lung cancer associ
29	184.5	18.2	273	21	AAW11680	Human secreted pro
30	184.5	18.2	273	22	ABB11989	Human secreted pro
31	184.5	18.2	273	22	AAW35809	Human protein sequ
32	184.5	18.2	273	22	AAW65619	Human colon cancer
33	183.5	18.1	225	22	AAW65272	Human cell surface
34	181	17.8	226	22	AAW98706	Chandra, a helper
35	180.5	17.8	227	22	ABG19236	Novel human diagno
36	176.5	17.4	267	20	AAW15224	Human receptor pro
37	176.5	17.4	273	22	ABG17004	Novel human diagno
38	174.5	17.2	299	21	AAW91352	Human secreted pro
39	172	16.9	178	22	AAE12073	Dendritic cell (DC
40	170	16.7	297	13	AAW20808	Human CD20 antigen
41	169	16.7	248	20	AAW36064	Extended human sec
42	168	16.6	297	10	AAW91356	CD20.4 antigen.
43	168	16.6	297	21	AAW96131	Human cell surface
44	168	16.6	297	22	AAU02440	Human lymphocyte c
45	167	16.5	297	17	AAW91436	Human CD20 antigen

ALIGNMENTS

RESULT 1	ID	AA001210	standard; Protein; 199 AA.
XX	AC	AA001210;	
XX	DT	26-SEP-2001	(first entry)
XX	DE	Novel human membrane protein #1.	
XX	KW	Human; membrane protein; membrane receptor	
XX	KW	physiological disorder.	
XX	OS	Homo sapiens.	
XX	FH	Key	Location/Qualifiers
FT	Misc-difference	199	/note= "Encoded by TGTGTG"
FT			
XX	PN	WO200146417-A2.	
XX	PD	28-JUN-2001.	
XX	PF	12-DEC-2000; 2000MO-US33742.	
XX	PR	22-DEC-1999; 99US-0171567.	
XX	PA	(LEXI-) LEXICON GENETICS INC.	
XX	PI	Walke DW, Turner CA;	
XX	WP1	2001-408646/43.	
DR	N-PSDB	AAS04279.	

XX Polynucleotide encoding novel human membrane protein, useful for
PT identifying agonist, antagonist or modifiers or for producing
PT antibodies useful in therapeutic, diagnostic and pharmacogenomic
PT applications
XX
XX Claim 2; Page 29-30; 32pp; English.
XX
XX The present sequence represents novel human membrane protein #1.
CC Human membrane protein #1 is 1 of 4 human membrane proteins
CC (AAU01210-AAU01213) given in the present invention. These membrane
CC proteins share structural similarity with membrane receptors such as
CC the 19E receptor and mammalian CD20. The novel human membrane proteins
CC are useful for identifying agonists, antagonists and modulators of the
CC membrane proteins. The membrane proteins can be used for diagnosis,
CC drug screening, pharmacogenomic applications, clinical trial monitoring
CC and the treatment of physiological disorders and diseases. The
CC polynucleotides encoding the membrane proteins can be used to generate
CC PCR primers or probes to identify mutations associated with a particular
CC disease.
XX
XX Sequence 199 AA:
SQ
Query Match 100.0%; Score 1015; DB 22; Length 199;
Best Local Similarity 100.0%; Pred. No. 3.2e-112;
Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 MDSSTASPVFLVPPPEITASEYESTELSATFTSTQSPLOKLFARKKILGTIQLFGIM 60
DB 1 MDSSTASPVFLVPPPEITASEYESTELSATFTSTQSPLOKLFARKKILGTIQLFGIM 60
OY 61 TFSGVIFLFTLLKPYRPPPIFLSGYPFGSVLFINSGAFLAVKRRKTEETLILSRIM 120
DB 61 TFSGVIFLFTLLKPYRPPPIFLSGYPFGSVLFINSGAFLAVKRRKTEETLILSRIM 120
OY 121 NLSALRAIGIITLTFEGFLDQNYICGYSHONSQCKAVTVLFGILITLMTFSIIELEFI 180
DB 121 NLSALRAIGIITLTFEGFLDQNYICGYSHONSQCKAVTVLFGILITLMTFSIIELEFI 180
OY 181 SLPFSILGCHSEDCCEOC 199
DB 181 SLPFSILGCHSEDCCEOC 199
RESULT 2
AAE13062
ID AAE13062 standard; Protein; 200 AA.
AC AAE13062;
XX
XX 28-JAN-2002 (first entry)
DE Human CD20/19E-receptor like protein, agp-96614-a1.
XX
XX Human; CD20/19E-receptor like protein; immunoglobulin E; agp-96614-a1;
KW agp-69406-a1; cancer; abnormal cell proliferation; autoimmune disease;
KW ovarian cancer; brain cancer; arteriosclerosis; vascular stenosis;
KW rheumatoid arthritis; multiple sclerosis; allergy; dermatitis; asthma;
KW reproductive disease; diabetes; transplant rejection; endometriosis;
KW infertility; gene therapy.
XX
XX Homo sapiens.
OS
XX
XX WO200174903-A2.
PN
XX 11-OCT-2001.
PD
XX
XX 29-MAR-2001; 2001WO-US10048.
PF
XX
XX 30-MAR-2000; 2000US-193728P.
PR
XX 27-NOV-2000; 2000US-0723258.
XX

PA (AMGE-) AMGEN INC.
XX
XX Welcher AA, Calzone FJ;
XX
XX WPI: 2001-662968/76.
DR N-PSDB; AAD21441.
DR
XX
XX Novel CD20/19E-receptor like polypeptides and polynucleotides,
PT antagonists and antibodies of the polypeptide useful for treating
PT ameliorating or preventing diseases associated with the polypeptide
PT e.g. cancer, asthma
XX
XX Claim 13; Fig 1; 145pp; English.
XX
XX The invention relates to human CD20/immunoglobulin E (19E)-receptor
CC like polypeptides designated as agp-96614-a1 and agp-69406-a1 and
CC nucleic acid molecules encoding such polypeptides. Polypeptides of
CC the invention are useful for treating, preventing or ameliorating
CC a disease, condition, or disorder which includes cancer such as
CC brain cancer, ovarian cancer; abnormal cell proliferation such as
CC arteriosclerosis, vascular stenosis; pathology from allergens
CC such as allergies, asthma, dermatitis; dysfunction of immune system
CC such as rheumatoid arthritis, autoimmune disease, multiple sclerosis,
CC diabetes; transplant rejection and reproductive diseases such as
CC infertility, preterm labour and delivery, endometriosis etc. They
CC are also useful for identifying antagonists and as immunogens, for
CC raising antibodies which may also be used to prevent, treat or
CC diagnose a number of diseases and disorders. Polynucleotides of the
CC invention are used to map the location of CD20/19E-receptor like
CC gene and related genes on chromosomes and as hybridisation probes.
CC They are also useful in gene therapy. The present sequence is
CC human CD20/19E-receptor like protein, agp-96614-a1.
XX
XX Sequence 200 AA:
SQ
Query Match 98.9%; Score 1004; DB 22; Length 200;
Best Local Similarity 99.0%; Pred. No. 6.6e-111;
Matches 197; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 MDSSTASPVFLVPPPEITASEYESTELSATFTSTQSPLOKLFARKKILGTIQLFGIM 60
DB 1 MDSSTASPVFLVPPPEITASEYESTELSATFTSTQSPLOKLFARKKILGTIQLFGIM 60
OY 61 TFSGVIFLFTLLKPYRPPPIFLSGYPFGSVLFINSGAFLAVKRRKTEETLILSRIM 120
DB 61 TFSGVIFLFTLLKPYRPPPIFLSGYPFGSVLFINSGAFLAVKRRKTEETLILSRIM 120
OY 121 NLSALRAIGIITLTFEGFLDQNYICGYSHONSQCKAVTVLFGILITLMTFSIIELEFI 180
DB 121 NLSALRAIGIITLTFEGFLDQNYICGYSHONSQCKAVTVLFGILITLMTFSIIELEFI 180
OY 181 SLPFSILGCHSEDCCEOC 199
DB 181 SLPFSILGCHSEDCCEOC 199
RESULT 3
AAG89138
ID AAG89138 standard; Protein; 200 AA.
AC AAG89138;
XX
XX 11-SEP-2001 (first entry)
DE Human secreted protein, SEQ ID NO: 258.
XX
XX Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;
KW GENSET.
XX
XX Homo sapiens.
OS
XX
XX WO200142451-A2.
PN
XX

PD 14-JUN-2001.
XX 07-DEC-2000; 2000MO-IB01938.
PF 08-DEC-1999; 99US-0169629.
XX 06-MAR-2000; 2000US-0187470.
XX (GEST) GENSET.
XX Dumas Malne Edwards J, Bougueleret L, Jobert S;
PI WPI; 2001-367870/38.
DR N-PSDB; AAH64741.
XX
XX Full length GENSET human nucleic acids encoding potentially secreted
PT proteins, useful in gene therapy and vaccination against a variety of
PT diseases, and for diagnosis of those diseases -
XX
XX Claim 21; Page 802-803; 921pp; English.
XX
XX The invention relates to full length GENSET human nucleic acids encoding
CC potentially secreted proteins. The nucleic acids and the polypeptides
CC they encode may be used in the prevention, treatment and diagnosis of
CC diseases associated with inappropriate GENSET gene expression. For
CC example, they be used to treat disorders associated with decreased
CC GENSET gene expression by rectifying mutations or deletions in a
CC patient's genome that affect the activity of GENSET or by supplementing
CC the patient's own production of GENSET polypeptides. Conversely,
CC antisense nucleic acid molecules may be administered to down regulate
CC GENSET expression by binding with the cells' own genes and preventing
CC their expression. The sense and antisense nucleic acids may also be
CC used as DNA probes in diagnostic assays to detect and quantitate the
CC presence of similar nucleic acid sequences in samples, and hence to
CC determine which patients may be in need of restorative therapy.
CC The GENSET polypeptides may be used as antigens in the production of
CC antibodies and in assays to identify modulators (agonists and
CC antagonists) of GENSET polypeptide expression and activity. The
CC present sequence is a GENSET polypeptide of the invention.
XX
XX
SO Sequence 200 AA;
Query Match 98.9%; Score 1004; DB 22; Length 200;
Best Local Similarity 99.0%; Pred. No. 6,6e-111;
Matches 197; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MOSSTASHPYFLVPEPEITASEYESTELSTSTSTQSPLOKLFARKMKILGTIQLFGIM 60
DB 1 MOSSTASHPYFLVPEPEITASEYESTELSTSTSTQSPLOKLFARKMKILGTIQLFGIM 60
QY 61 TFSFGVIFLEFTLLKPYRPFPIFLSGYPFWGVLFTNSGAFLLAVRKKTETLLISRIIM 120
DB 61 TFSFGVIFLEFTLLKPYRPFPIFLSGYPFWGVLFTNSGAFLLAVRKKTETLLISRIIM 120
QY 121 NLSALRAIAGIILLTFGFLIDQNYICGYSHQNSOCKAVVLEFLIGLITLMTFSIIELEF 180
DB 121 NLSALRAIAGIILLTFGFLIDQNYICGYSHQNSOCKAVVLEFLIGLITLMTFSIIELEF 180
QY 181 SLFPSILGCHSEDCDEOC 199
DB 181 SLFPSILGCHSEDCDEOC 199
RESULT 4
AAE10917
ID AAE10917 standard; Protein; 201 AA.
XX
XX AAE10917;
XX
XX 18-DEC-2001 (first entry)
XX
XX Human gene 9 encoded immune system-related protein HTENNA5.
DE
XX Human: immune system-related protein; allergy; rheumatoid arthritis;
KW

KW cancer; Gaucher's disease; viral hepatitis; gene therapy; cytostatic;
KW diabetes mellitus; arhythmia; wound healing; ischaemic lesion; AIDS;
KW Acquired Immune Deficiency Syndrome; vitruide; hepatotropic; vasotropic;
KW autoimmune disorder; inflammation; cardiovascular disorder; hair loss;
KW wound healing; cell proliferation; skin aging; endocrine disorder;
XX food preservative.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH Domain 21..26
FT Domain /label= Immunogenic_epitope
FT Domain 150..156
FT Domain /label= Immunogenic_epitope
PN WO200166722-A1.
XX
XX 13-SEP-2001.
PD
PF 07-MAR-2001; 2001MO-US07260.
XX
XX 08-MAR-2000; 2000US-187873P.
PR 11-AUG-2000; 2000US-224367P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ni J, Hilbert D, Kenny JJ, Moore PA, Choi GH, Soppet DR, Ebner R;
PI Gruber JR, Endress GA, Ruben SM;
DR N-PSDB; AAD18275.
XX
XX WPI: 2001-589939/66.
XX
XX Novel isolated immune system-related polypeptide useful for treating
PT rheumatoid arthritis, AIDS, allergy, cancer, Gaucher's disease,
PT diabetes mellitus, arrhythmia, wound healing, ischemic lesions and
PT viral hepatitis -
XX
XX Claim 11; Page 310-311; 315pp; English.
XX
XX The invention relates to human immune system-related protein and their
CC DNA. Human immune-system related protein and DNA are useful for
CC preventing, treating or ameliorating a medical condition in a mammalian
CC subject, for diagnosing, preventing or treating immune system-associated
CC disorders, autoimmune disorders (rheumatoid arthritis), inflammatory
CC disorders (allergies), immunodeficiencies (AIDS), inflammatory disorders
CC (allergies), infectious diseases (e.g., viral hepatitis), complement
CC activation disorders, immune complex diseases, neoplastic disorders
CC (cancer), hyperproliferative disorders (Gaucher's disease), disorders
CC associated with neovascularisation, diseases at the cellular level,
CC cardiovascular disorders (arrhythmias), wound healing and epithelial
CC cell proliferation, endocrine disorders (diabetes mellitus) and
CC neurological disorders (ischaemic lesions). Immune-system related protein
CC or DNA is useful for preventing hair loss, skin aging due to sunburn, to
CC maintain organs before transplantation, to treat weight disorders, to
CC modulate mammalian characteristics, to change a mammal's mental or
CC physical state, or as a food additive or preservative. Immune-system
CC related DNA is useful in gene therapy, for chromosome identification,
CC radiation hybrid mapping, long range restriction mapping and in forensic
CC biology. The present sequence represents a human immune-system related
CC protein of the invention.
XX
XX
SO Sequence 201 AA;
Query Match 90.9%; Score 923; DB 22; Length 201;
Best Local Similarity 98.4%; Pred. No. 2,8e-101;
Matches 185; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 MOSSTASHPYFLVPEPEITASEYESTELSTSTSTQSPLOKLFARKMKILGTIQLFGIM 60
DB 1 MOSSTASHPYFLVPEPEITASEYESTELSTSTSTQSPLOKLFARKMKILGTIQLFGIM 60
QY 61 TFSFGVIFLEFTLLKPYRPFPIFLSGYPFWGVLFTNSGAFLLAVRKKTETLLISRIIM 120
DB 61 TFSFGVIFLEFTLLKPYRPFPIFLSGYPFWGVLFTNSGAFLLAVRKKTETLLISRIIM 120

```
Db      61 TEFSGVIFLTLPKPYRPFIFLSCGYPWGSVLFINSAGFLIAVRKKTETIIILSRIM 120
QY      121 NLLSALRAAGIILTFEFGILLDQNYICGYSHQNSOCKAVTVLFIIGILITLMPFSIIELEFI 180
Db      121 NPLSALGAIGIILTFEFGILLDQNYICGYSHQNSOCKAVTVLFIIGILITLMPFSIIELEFI 180
QY      181 SLFPSILG 188
Db      181 SLFPSIWG 188

RESULT 5
ABR95818
ID      ABR95818 standard; Protein; 212 AA.
XX
AC      ABR95818;
XX
DT      21-JUN-2002 (first entry)
DE      Human testicular antigen SEQ ID NO: 1202.
XX
KW      Human; testicular antigen; testes; cancer; metastasis; immune disorder;
KW      reproductive system disorder; urinary system disorder; gene therapy;
KW      cardiovascular disorder; respiratory disorder; neurological disorder;
KW      gastrointestinal disease; infection; cytostatic.
OS      Homo sapiens.
XX
PN      WO200155317-A2.
XX
PD      02-AUG-2001.
XX
PF      17-JAN-2001; 2001WO-US01329.
XX
PR      31-JAN-2000; 2000US-0179065.
PR      04-FEB-2000; 2000US-0180628.
PR      24-FEB-2000; 2000US-0184664.
PR      02-MAR-2000; 2000US-0186350.
PR      16-MAR-2000; 2000US-0189874.
PR      17-MAR-2000; 2000US-019076.
PR      18-APR-2000; 2000US-0198123.
PR      19-MAY-2000; 2000US-0205515.
PR      07-JUN-2000; 2000US-0209467.
PR      28-JUN-2000; 2000US-0214886.
PR      30-JUN-2000; 2000US-0215135.
PR      07-JUL-2000; 2000US-0216647.
PR      07-JUL-2000; 2000US-0216880.
PR      11-JUL-2000; 2000US-0217487.
PR      11-JUL-2000; 2000US-0217496.
PR      14-JUL-2000; 2000US-0218290.
PR      26-JUL-2000; 2000US-0220963.
PR      14-AUG-2000; 2000US-0220964.
PR      14-AUG-2000; 2000US-0224518.
PR      14-AUG-2000; 2000US-0224519.
PR      14-AUG-2000; 2000US-0225213.
PR      14-AUG-2000; 2000US-0225214.
PR      14-AUG-2000; 2000US-0225266.
PR      14-AUG-2000; 2000US-0225267.
PR      14-AUG-2000; 2000US-0225268.
PR      14-AUG-2000; 2000US-0225270.
PR      14-AUG-2000; 2000US-0225447.
PR      14-AUG-2000; 2000US-0225757.
PR      14-AUG-2000; 2000US-0225758.
PR      14-AUG-2000; 2000US-0225759.
PR      18-AUG-2000; 2000US-0225729.
PR      22-AUG-2000; 2000US-0226881.
PR      22-AUG-2000; 2000US-0226868.
PR      22-AUG-2000; 2000US-0227182.
PR      23-AUG-2000; 2000US-0227009.
PR      30-AUG-2000; 2000US-0228924.
PR      01-SEP-2000; 2000US-0229287.
PR      01-SEP-2000; 2000US-0229343.
PR      01-SEP-2000; 2000US-0229344.

PR      01-SEP-2000; 2000US-0229345.
PR      05-SEP-2000; 2000US-0229509.
PR      05-SEP-2000; 2000US-0229513.
PR      06-SEP-2000; 2000US-0230437.
PR      06-SEP-2000; 2000US-0230438.
PR      08-SEP-2000; 2000US-0231242.
PR      08-SEP-2000; 2000US-0231243.
PR      08-SEP-2000; 2000US-0231244.
PR      08-SEP-2000; 2000US-0231413.
PR      08-SEP-2000; 2000US-0231414.
PR      08-SEP-2000; 2000US-0232080.
PR      08-SEP-2000; 2000US-0232081.
PR      12-SEP-2000; 2000US-0231968.
PR      14-SEP-2000; 2000US-0233397.
PR      14-SEP-2000; 2000US-0233398.
PR      14-SEP-2000; 2000US-0233399.
PR      14-SEP-2000; 2000US-0233400.
PR      14-SEP-2000; 2000US-0233401.
PR      14-SEP-2000; 2000US-0233063.
PR      14-SEP-2000; 2000US-0233064.
PR      14-SEP-2000; 2000US-0233065.
PR      21-SEP-2000; 2000US-0234223.
PR      21-SEP-2000; 2000US-0234274.
PR      25-SEP-2000; 2000US-0234997.
PR      25-SEP-2000; 2000US-0234998.
PR      26-SEP-2000; 2000US-0235484.
PR      27-SEP-2000; 2000US-0235834.
PR      27-SEP-2000; 2000US-0235836.
PR      29-SEP-2000; 2000US-0236327.
PR      29-SEP-2000; 2000US-0236367.
PR      29-SEP-2000; 2000US-0236368.
PR      29-SEP-2000; 2000US-0236369.
PR      29-SEP-2000; 2000US-0236370.
PR      02-OCT-2000; 2000US-0236802.
PR      02-OCT-2000; 2000US-0237037.
PR      02-OCT-2000; 2000US-0237038.
PR      02-OCT-2000; 2000US-0237039.
PR      13-OCT-2000; 2000US-0237040.
PR      13-OCT-2000; 2000US-0239335.
PR      13-OCT-2000; 2000US-0239337.
PR      20-OCT-2000; 2000US-0240960.
PR      20-OCT-2000; 2000US-0241221.
PR      20-OCT-2000; 2000US-0241785.
PR      20-OCT-2000; 2000US-0241786.
PR      20-OCT-2000; 2000US-0241787.
PR      20-OCT-2000; 2000US-0241808.
PR      20-OCT-2000; 2000US-0241809.
PR      20-OCT-2000; 2000US-0241826.
PR      01-NOV-2000; 2000US-0244617.
PR      08-NOV-2000; 2000US-0246474.
PR      08-NOV-2000; 2000US-0246475.
PR      08-NOV-2000; 2000US-0246476.
PR      08-NOV-2000; 2000US-0246477.
PR      08-NOV-2000; 2000US-0246478.
PR      08-NOV-2000; 2000US-0246523.
PR      08-NOV-2000; 2000US-0246524.
PR      08-NOV-2000; 2000US-0246525.
PR      08-NOV-2000; 2000US-0246526.
PR      08-NOV-2000; 2000US-0246527.
PR      08-NOV-2000; 2000US-0246528.
PR      08-NOV-2000; 2000US-0246532.
PR      08-NOV-2000; 2000US-0246609.
PR      08-NOV-2000; 2000US-0246610.
PR      08-NOV-2000; 2000US-0246611.
PR      08-NOV-2000; 2000US-0246613.
PR      17-NOV-2000; 2000US-0249207.
PR      17-NOV-2000; 2000US-0249208.
PR      17-NOV-2000; 2000US-0249209.
PR      17-NOV-2000; 2000US-0249210.
PR      17-NOV-2000; 2000US-0249211.
PR      17-NOV-2000; 2000US-0249212.
PR      17-NOV-2000; 2000US-0249213.
PR      17-NOV-2000; 2000US-0249214.
```


[illegible]

PR	08-DEC-2000;	2000US-0251869.	
PR	08-DEC-2000;	2000US-0251989.	
PR	08-DEC-2000;	2000US-0251990.	
PR	11-DEC-2000;	2000US-0254097.	
PR	05-JAN-2001;	2001US-0259678.	
XX			
PA	(HUMA-)	HUMAN GENOME SCL INC.	
XX			
PI	Rosen CA,	Barash SC, Ruben SM;	
XX			
DR	WPI: 2001-465570/50.		
DR	N-PSDB: AAL01084.		
XX			
PT	Isolated nucleic acid molecule encoding a reproductive system antigen		
PR	is used in preventing, treating or ameliorating a medical condition -		
XX			
PS	Claim 11; SEQ ID NO 3772; 1297bp + Sequence Listing; English.		
XX			
CC	The present invention provides the protein and coding sequences of a		
CC	number of human reproductive system related antigens. These can be used		
CC	in the prevention and treatment of reproductive system disorders,		
CC	including cancer. The present sequence is a protein of the invention.		
XX			
SO	Sequence	212 AA;	
	Query Match	90.9%; Score 923; DB 22; Length 212;	
	Best Local Similarity	98.4%; Pred. No. 3e-101;	
	Matches	185; Conservative	0; Mismatches
			3; Indels
			0; Gaps
OY	1	MDSSTASHSPVLVPPPEPTASGETESTELSATFTFSQSPLOKLFARKMKILGTIQLFGIM	60
DB	12	MDSSTASHSPVLFPPPEPTASEESTELSATFTFSQSPLOKLFARKMKILGTIQLFGIM	71
OY	61	TFSEVFILFTLKRPYPRPFTELSGYPFGSVLFINSGAFLIAVKRKTEETLIISRIM	120
DB	72	TFSEVFILFTLKRPYPRPFTELSGYPFGSVLFINSGAFLIAVKRKTEETLIISRIM	131
OY	121	NLSLSRAIRIAGIILITFGFILDQNYICGSHQNSOCKKATVLFGLITLTMTFSITELFI	180
DB	132	NLSLSRAIRIAGIILITFGFILDQNYICGSHQNSOCKKATVLFGLITLTMTFSITELFI	191
OY	181	SLPFSITLG	188
DB	192	SLPFSITWG	199
	RESULT 7		
	AA001213		
ID	AA001213	standard; Protein; 149 AA.	
XX			
AC	AA001213;		
XX			
DT	26-SEP-2001	(first entry)	
XX			
DE	Novel human membrane protein #4.		
XX			
KW	Human; membrane protein; membrane receptor; Ige receptor; CD20;		
KW	physiological disorder.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200146417-A2.		
XX			
PD	28-JUN-2001.		
XX			
PF	12-DEC-2000; 2000WO-US33742.		
XX			
PR	22-DEC-1999; 99US-0171567.		
XX			
EA	(LEXI-) LEXICON GENETICS INC.		
XX			
PI	Walke DW, Turner CA;		
XX			

DR WPI: 2001-408646/43.
DR N-PSDB: AAS04282.
XX
PT Polynucleotide encoding novel human membrane protein, useful for
PT identifying agonist, antagonist or modifiers or for producing
PT antibodies useful in therapeutic, diagnostic and pharmacogenomic
PT applications -
XX
PS Claim 4: Page 31; 32pp; English.
XX
CC The present sequence represents novel human membrane protein #4.
CC Human membrane protein #4 is 1 of 4 human membrane proteins
CC (AAU01210-AAU01213) given in the present invention. These membrane
CC proteins share structural similarity with membrane receptors such as
CC the Ige receptor and mammalian CD20. The novel human membrane proteins
CC are useful for identifying agonists, antagonists and modulators of the
CC membrane proteins, and for producing antibodies specific to the
CC membrane proteins. The membrane proteins can be used for diagnosis,
CC drug screening, pharmacogenomic applications, clinical trial monitoring
CC and the treatment of physiological disorders and diseases. The
CC polynucleotides encoding the membrane proteins can be used to generate
CC PCR primers or probes to identify mutations associated with a particular
CC disease.
XX
SQ Sequence 149 AA:
XX
Query Match 71.6%; Score 726.5; DB 22; Length 149;
Best Local Similarity 74.4%; Pred. No. 4,2e-78;
Matches 148; Conservative 0; Mismatches 0; Indels 51; Gaps 1;
QY 1 MDSSTASHPVFLVFPPEITASESESTELSATPTSPLOKLPARKKILGTIQLIFGIM 60
DB 1 MDSSTASHPVFLVFPPEITASESESTELSATPTSPLOKLPARKKILGTIQLIFGIM 60
QY 61 TFSFGVIFLFTLLKPYRPPPIFLSGYPWGSVLFINSGLAVARKTETLIILSRIM 120
DB 61 TFSFGVIFLFTLLKPYRPPPIFLSGYPWGSVLFINSGLAVARKTETLIILSRIM 120
QY 121 NLSALRAINAGILILTFEGLIDQNYICGYSHNSOCKAVTLFLGLITLMTFSITELFI 180
DB 113 -----LGLITLMTFSITELFI 129
QY 181 SLFPSILGCHSEDCDEQC 199
DB 130 SLFPSILGCHSEDCDEQC 148
RESULT 8
ABBI2234
ID ABBI2234 standard; peptide; 158 AA.
XX
AC ABBI2234;
XX
DE Human secreted protein homologue, SEQ ID NO:2604.
XX
DT 11-JAN-2002 (first entry)
XX
XX Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;
KW bone disorder; osteoporosis; vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;
KW cell culture; drug screening; gene therapy; antiinflammatory;
KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW antifungal; vulnery; antilucer.
XX
XX Homo sapiens.
XX
OS

PN MO200157188-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US03800.
XX
PR 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
XX
XX (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
DR WPI: 2001-457740/49.
DR N-PSDB: ABA09478.
XX
PT Human proteins and DNA encoding sequences useful for preventing,
PT treating or ameliorating a medical condition in a mammalian subject
PT e.g. arthritis and cancer -
XX
PS Claim 20: Page 318; 1963pp; English.
XX
CC Sequences ABBI0981-ABBI2330 represent 1350 novel human polypeptides, and
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC invention also relates to vectors and recombinant host cells comprising a
CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotides
CC or polypeptides in a sample, and methods of identifying compounds which
CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, thereby
CC giving an insight into their probable biological activities, and hence
CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation activities; stem cell growth factor activity;
CC haematopoiesis regulatory activity; tissue growth activity;
CC immunomodulatory activity; activin- or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be
CC involved in oncogenesis, cancer cell proliferation or metastasis.
CC Depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.
CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a novel human
CC polypeptide of the invention.
XX
SQ Sequence 158 AA:
XX
Query Match 66.3%; Score 673; DB 22; Length 158;
Best Local Similarity 97.8%; Pred. No. 1e-71;
Matches 136; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 MDSSTASHPVFLVFPPEITASESESTELSATPTSPLOKLPARKKILGTIQLIFGIM 60
DB 12 MDSSTASHPVFLVFPPEITASESESTELSATPTSPLOKLPARKKILGTIQLIFGIM 71
QY 61 TFSFGVIFLFTLLKPYRPPPIFLSGYPWGSVLFINSGLAVARKTETLIILSRIM 120
DB 72 TFSFGVIFLFTLLKPYRPPPIFLSGYPWGSVLFINSGLAVARKTETLIILSRIM 131

QY 121 NLSALRAIAGIILLTFGE 139
| | | | | | | | | | | | | | | |
Db 132 NLSALGAIAGIILLTFEE 150

RESULT 9

AAU01212
ID AAU01212 standard; Protein: 138 AA.

AC AAU01212;

DT 26-SEP-2001 (first entry)

DE Novel human membrane protein #3.

Human; membrane protein; membrane receptor; Ige receptor; CD20; physiological disorder.

OS Homo sapiens.

PN WO200146417-A2.

PD 28-JUN-2001.

PF 12-DEC-2000; 2000WO-US33742.

PR 22-DEC-1999; 99US-0171567.

PA (LEXI-) LEXICON GENETICS INC.

PI Walke DW, Turner CA;

DR WPI; 2001-408646/43.

N-PSDB; AAS04281.

PT Polynucleotide encoding novel human membrane protein, useful for identifying agonist, antagonist or modifiers or for producing antibodies useful in therapeutic, diagnostic and pharmacogenomic applications

PS Disclosure; Page 30-31; 32pp; English.

CC The present sequence represents novel human membrane protein #3.

CC Human membrane protein #3 is 1 of 4 human membrane proteins

CC (AAU01210-AAU01213) given in the present invention. These membrane

CC proteins share structural similarity with membrane receptors such as

CC the Ige receptor and mammalian CD20. The novel human membrane proteins

CC are useful for identifying agonists, antagonists and modulators of the

CC membrane proteins, and for producing antibodies specific to the

CC drug screening, pharmacogenomic applications, clinical trial monitoring

CC and the treatment of physiological disorders and diseases. The

CC polynucleotides encoding the membrane proteins can be used to generate

CC PCR primers or probes to identify mutations associated with a particular

CC disease.

CC Sequence 138 AA;

QQ

Query Match

Best Local Similarity 57.4%; Score 583; DB 22; Length 138;

Matches 117; Conservative 5; Mismatches 0; Indels 6; Gaps 1;

QY 1 MDSSTAHSPVLPPEITASEYESTELSATFTSTQSPLOKLFARKMKILGTIOILFGIM 60

Db 1 MDSSTAHSPVLPPEITASEYESTELSATFTSTQSPLOKLFARKMKILGTIOILFGIM 60

QY 61 TFSFGVIFLTLKRPYPRPFIFLSGYPWGVLFINSAGFLIAVKRKTTETLI----- 114

Db 61 TFSFGVIFLTLKRPYPRPFIFLSGYPWGVLFINSAGFLIAVKRKTTETLI----- 114

QY 115 ILSRTIMNL 122

Db 115 ILSRTIMNL 122

Db 121 ILTKIVSV 128

RESULT 10

AAU01211
ID AAU01211 standard; Protein: 81 AA.

AC AAU01211;

DT 26-SEP-2001 (first entry)

DE Novel human membrane protein #2.

Human; membrane protein; membrane receptor; Ige receptor; CD20; physiological disorder.

OS Homo sapiens.

PN WO200146417-A2.

PD 28-JUN-2001.

PF 12-DEC-2000; 2000WO-US33742.

PR 22-DEC-1999; 99US-0171567.

PA (LEXI-) LEXICON GENETICS INC.

PI Walke DW, Turner CA;

DR WPI; 2001-408646/43.

N-PSDB; AAS04280.

PT Polynucleotide encoding novel human membrane protein, useful for identifying agonist, antagonist or modifiers or for producing antibodies useful in therapeutic, diagnostic and pharmacogenomic applications

PS Disclosure; Page 30; 32pp; English.

CC The present sequence represents novel human membrane protein #2.

CC Human membrane protein #2 is 1 of 4 human membrane proteins

CC (AAU01210-AAU01213) given in the present invention. These membrane

CC proteins share structural similarity with membrane receptors such as

CC the Ige receptor and mammalian CD20. The novel human membrane proteins

CC are useful for identifying agonists, antagonists and modulators of the

CC membrane proteins, and for producing antibodies specific to the

CC drug screening, pharmacogenomic applications, clinical trial monitoring

CC and the treatment of physiological disorders and diseases. The

CC polynucleotides encoding the membrane proteins can be used to generate

CC PCR primers or probes to identify mutations associated with a particular

CC disease.

CC Sequence 81 AA;

QQ

Query Match 41.0%; Score 416; DB 22; Length 81;

Best Local Similarity 100.0%; Pred. No. 1.5e-41;

Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 NMLSLALRAIAGIILLTFGFILIDQNYICGSHNSQCAVYVLFGLITLTMTFSIIEF 179

Db 1 NMLSLALRAIAGIILLTFGFILIDQNYICGSHNSQCAVYVLFGLITLTMTFSIIEF 179

QY 180 ISLPSIIIGCHSEDCDCQC 199

Db 61 ISLPSIIIGCHSEDCDCQC 199

QY 61 ISLPSIIIGCHSEDCDCQC 199

Db 61 ISLPSIIIGCHSEDCDCQC 199

QY 115 ILSRTIMNL 122

Db 115 ILSRTIMNL 122

QY 115 ILSRTIMNL 122

Db 115 ILSRTIMNL 122

QY 115 ILSRTIMNL 122

Db 115 ILSRTIMNL 122

QY 115 ILSRTIMNL 122

Db 115 ILSRTIMNL 122

QY 115 ILSRTIMNL 122

Db 115 ILSRTIMNL 122

AC AAG03074;
XX
XX 06-OCT-2000 (first entry)
XX
DE Human secreted protein, SEQ ID NO: 7155.
XX
XX Human, 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping.
XX
XX Homo sapiens.
OS
XX EPI033401-A2.
PN
XX 06-SEP-2000.
PD
XX 21-FEB-2000; 2000EP-0200610.
PE
XX 26-FEB-1999; 990S-0122487.
PR
XX (GENSET) GENSET.
PA
XX Dumas Milne Edwards J, Duclert A, Giordano J;
PI WPI: 2000-500381/45.
DR N-PSDB; AAC03080.
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
XX Claim 13; SEQ ID 7155; 71pp + CD-ROM; English.
PS
XX The present sequence is a polypeptide encoded by one of a large number
CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
CC were prepared from total human RNAs or polyA+ RNAs derived from 30
CC different tissues. EST sequences usually correspond mainly to the 3'
CC untranslated region (UTR) of the mRNA because they are often obtained
CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
CC those cases where longer cDNA sequences have been obtained, the full 5'
CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
CC ends and can therefore be used to obtain full length cDNAs and genomic
CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
CC chromosome mapping procedures. They are used to obtain upstream
CC regulatory sequences and to design expression and secretion vectors.
XX
XX Sequence 77 AA:
SQ
Query Match 29.0%; Score 294; DB 21; Length 77;
Best Local Similarity 83.6%; Pred. No. 4,4e-27;
Matches 56; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
QY 33 ESTQSPQLQFLARKKMLIGTQILFGIMTFSGVIFLFTLLKPYRPFPIFLSGYPMGQS 92
DB 11 FOLKAPCKLFLARKKMLIGTQILFGIMTFSGVIFLFTLLKPYRPFPIFLSGYPMGQS 70
QY 93 VLFINSG 99
DB 71 VLFINSG 77
RESULT 12
AAG89142
ID AAG89142 standard; Protein; 67 AA.
XX
XX AAG89142;
AC
XX
XX 11-SEP-2001 (first entry)
DE Human secreted protein, SEQ ID NO: 262.
XX
XX Human, secreted protein; gene therapy; vaccine; treatment; diagnosis;
KW GENSET.

XX
XX Homo sapiens.
OS
XX WO200142451-A2.
PN
XX 14-JUN-2001.
PD
XX 07-DEC-2000; 2000WO-IB01938.
PE
XX 08-DEC-1999; 990S-0169629.
PR 06-MAR-2000; 2000US-0187470.
XX
XX (GENSET) GENSET.
PA
XX Dumas Milne Edwards J, Bougueleret L, Jobert S;
PI WPI: 2001-367870/38.
DR N-PSDB; AAH64745.
XX
XX Full length GENSET human nucleic acids encoding potentially secreted
PT proteins, useful in gene therapy and vaccination against a variety of
PT diseases, and for diagnosis of those diseases -
XX
XX Claim 21; Page 805; 921pp; English.
PS
XX The invention relates to full length GENSET human nucleic acids encoding
CC potentially secreted proteins. The nucleic acids and the polypeptides
CC they encode may be used in the prevention, treatment and diagnosis of
CC diseases associated with inappropriate GENSET gene expression. For
CC example, they be used to treat disorders associated with decreased
CC GENSET gene expression by rectifying mutations or deletions in a
CC patient's genome that affect the activity of GENSET or by supplementing
CC the patients own production of GENSET polypeptides. Conversely,
CC antisense nucleic acid molecules may be administered to down regulate
CC GENSET expression by binding with the cells' own genes and preventing
CC their expression. The sense and antisense nucleic acids may also be
CC used as DNA probes in diagnostic assays to detect and quantitate the
CC presence of similar nucleic acid sequences in samples, and hence to
CC determine which patients may be in need of restorative therapy.
CC The GENSET polypeptides may be used as antigens in the production of
CC antibodies and in assays to identify modulators (agonists and
CC antagonists) of GENSET polypeptide expression and activity. The
CC present sequence is a GENSET polypeptide of the invention.
XX
XX Sequence 67 AA:
SQ
Query Match 26.0%; Score 263.5; DB 22; Length 67;
Best Local Similarity 74.7%; Pred. No. 1.5e-23;
Matches 56; Conservative 3; Mismatches 7; Indels 9; Gaps 1;
QY 1 MDSTAHSPVFLVFPPEITASSESTELSATTFSTQSPLOKFLARKKMLIGTQILFGIM 60
DB 1 MDSTAHSPVFLVFPPEITASSESTELSATTFSTQSPLOKFLARKKMLIGTQILFGIM 54
QY 61 TFSFGVIFLFTLLK 75
DB 55 ---SGALFCSLIIIEP 66
RESULT 13
ABB90317
ID ABB90317 standard; Protein; 220 AA.
XX
XX ABB90317;
AC
XX
XX 24-MAY-2002 (first entry)
DE Human polypeptide SEQ ID NO 2693.
XX
XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
KW vunerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;

AAAY50174

ID	AA	Y50174	standard; Protein; 239 AA.
1	Met	1	Met
2	Met	2	Met
3	Met	3	Met
4	Met	4	Met
5	Met	5	Met
6	Met	6	Met
7	Met	7	Met
8	Met	8	Met
9	Met	9	Met
10	Met	10	Met
11	Met	11	Met
12	Met	12	Met
13	Met	13	Met
14	Met	14	Met
15	Met	15	Met
16	Met	16	Met
17	Met	17	Met
18	Met	18	Met
19	Met	19	Met
20	Met	20	Met
21	Met	21	Met
22	Met	22	Met
23	Met	23	Met
24	Met	24	Met
25	Met	25	Met
26	Met	26	Met
27	Met	27	Met
28	Met	28	Met
29	Met	29	Met
30	Met	30	Met
31	Met	31	Met
32	Met	32	Met
33	Met	33	Met
34	Met	34	Met
35	Met	35	Met
36	Met	36	Met
37	Met	37	Met
38	Met	38	Met
39	Met	39	Met
40	Met	40	Met
41	Met	41	Met
42	Met	42	Met
43	Met	43	Met
44	Met	44	Met
45	Met	45	Met
46	Met	46	Met
47	Met	47	Met
48	Met	48	Met
49	Met	49	Met
50	Met	50	Met
51	Met	51	Met
52	Met	52	Met
53	Met	53	Met
54	Met	54	Met
55	Met	55	Met
56	Met	56	Met
57	Met	57	Met
58	Met	58	Met
59	Met	59	Met
60	Met	60	Met
61	Met	61	Met
62	Met	62	Met
63	Met	63	Met
64	Met	64	Met
65	Met	65	Met
66	Met	66	Met
67	Met	67	Met
68	Met	68	Met
69	Met	69	Met
70	Met	70	Met
71	Met	71	Met
72	Met	72	Met
73	Met	73	Met
74	Met	74	Met
75	Met	75	Met
76	Met	76	Met
77	Met	77	Met
78	Met	78	Met
79	Met	79	Met
80	Met	80	Met
81	Met	81	Met
82	Met	82	Met
83	Met	83	Met
84	Met	84	Met
85	Met	85	Met
86	Met	86	Met
87	Met	87	Met
88	Met	88	Met
89	Met	89	Met
90	Met	90	Met
91	Met	91	Met
92	Met	92	Met
93	Met	93	Met
94	Met	94	Met
95	Met	95	Met
96	Met	96	Met
97	Met	97	Met
98	Met	98	Met
99	Met	99	Met
100	Met	100	Met
101	Met	101	Met
102	Met	102	Met
103	Met	103	Met
104	Met	104	Met
105	Met	105	Met
106	Met	106	Met
107	Met	107	Met
108	Met	108	Met
109	Met	109	Met

AC AAY50174;

DT 31-JAN-2000 (first entry)

Human high affinity IgE receptor-like protein (IGERB).

KW IGFRRB: high affinity; immunoglobulin E; IGE; receptor; allergy;
KW inflammation: B lymphocyte; homology; high affinity IGE receptor; FCRI;
KW mast cell; basophil; histamine; protease; cytoplasmic granule; synthesis
KW effector; prostaglandin; leukotriene; cytokine; antigen; antibody;
KW identification; agonist; antagonists; expression; activity; diagnosis;
KW therapy; inflammatory disorder; multiple sclerosis; osteoarthritis;
KW asthma; cancer; side effect; complication.

OS Homo sapiens.

FH	Key	Location/Qualifiers
FH	Key	Location/Qualifiers

ET	Modified-site	125	/octo- [Phosphoryl] ated by metalic phosphate of
ET	Modified-site	125	/octo- [Phosphoryl] ated by metalic phosphate of

FT	Modified-site	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	539	540	541	542	543	544	545	546	547	548	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567	568	569	570	571	572	573	574	575	576	577	578	579	580	581	582	583	584	585	586	587	588
----	---------------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

Modified-site 192

PN US5977072-A

02-NOV-1999 PD

PF 15-DEC-1998; 98US-0213389.

PR 21-AUG-1997; 97US-0916902.

PA (INCY-) INCYTE PHARM INC.

PI Bandman O, Corley NC, Lal P, xv

DR WPI; 2000-012123/01.
DB N-DEDA 28732943

XX
XX
H
2
b
u
e
c
h
u
s
t
i
s
t
r
y

stimulating allergic and immune responses -

PS Claim 1; Fig 1; 29pp; English
 XY

This sequence represents human high affinity immunoglobulin E (IgE) receptor-like protein (IGERB). Nucleic acids encoding IGERB were initially identified in a brain cDNA library, this sequence being a consensus. An allergic response is initiated by release of IgE from B lymphocytes. The IgE molecules then bind to the high affinity IgE receptor (FcεR) present on mast cells and basophils, which triggers the release of histamine and proteases from cytoplasmic granules and leads to the synthesis of effectors of the allergic and inflammatory response, such as prostaglandins, leukotrienes and cytokines. As IGERB binds IgE, it may be administered to stimulate allergic and immune responses in patients in whom IGERB is under expressed or inactive and to supplement the patients own production of the protein. IGERB may also be used as an antigen for the production of antibodies and to identify candidate agonists and antagonists of IGERB expression and activity. The antibodies may also be used in diagnosis. Antibodies and antagonists may be administered to downregulate IGERB activity and reduce the potency of inflammatory and allergic responses. They may be used in this way to treat inflammatory disorders such as multiple sclerosis, osteoarthritis, asthma and some complications of cancer. Conversely, the agonists may be used to enhance immune responses.

sq	Sequence	239	AA
----	----------	-----	----

Query Match	20.3%;	Score 206.5;	DB 21;	Length 239;
Best Local Similarity	31.8%;	Pred. No. 4.8e-16;		
Matches 49;	Conservative 32;	Mismatches 70;	Indels 3;	Gaps 2

39 LQKLEAR-KMKILGTIQILFGIMTFSEGVIFLFTLLKPYRPFIFLSGYPEWGSVLEIN 97

Db 54 LQEKFLKGEPKVLGVQILTALMSLSMGITMCMASNTYGSNPISVYIGYTIWGSVMFI 113

98 SCAPLIAVKKRTETLILSRIMLLSARAIAGIILLTF--GFILDQNYICGYSHQNSQ 155

Db 1.14 SCSLSIAGIRTTKGLVRGSLGMNITSSVLAASGILINTFSLAFYFHHPPYCNYYGNSNN 173

QY 156 CKAATVLFGLITMTFSIIEFLISLPFSIIGC 189

Db 174 CHGTMSILMGLDGMVLLSVLEFCIAVSLSAFCG 207

Search completed: February 24, 2003, 13:02:53
Job time : 40.3132 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2003, 13:01:46 : Search time 22.8736 Seconds
(without alignments)
255.979 Million cell updates/sec

Title: US-09-735-712-2

Perfect score: 1015

Sequence: 1 MDSTAHSPVFLVFPPEITA.....ISLPSILGCHSEDCDCRC 199

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/CTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfilest1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	206.5	20.3	239	2	US-08-916-902A-1
2	206.5	20.3	239	2	US-09-213-389-1
3	184.5	18.2	214	1	US-08-318-492-4
4	184.5	18.2	214	1	US-08-707-340-4
5	184.5	18.2	214	2	US-08-916-902A-3
6	184.5	18.2	214	2	US-08-994-578-4
7	184.5	18.2	214	2	US-09-213-389-3
8	159	15.7	243	1	US-07-869-933-29
9	159	15.7	243	1	US-07-869-933-33
10	159	15.7	243	1	US-08-201-879A-4
11	159	15.7	243	2	US-08-916-902A-4
12	159	15.7	243	2	US-09-213-389-4
13	159	15.7	243	4	US-09-103-663-29
14	159	15.7	243	4	US-09-103-663-33
15	159	15.7	246	4	US-07-869-933-23
16	159	15.7	246	4	US-09-103-663-23
17	158.5	15.6	235	1	US-07-869-933-34
18	158.5	15.6	235	1	US-08-201-879A-5
19	158.5	15.6	235	4	US-09-103-663-34
20	158.5	15.6	247	4	US-09-724-864-49
21	123.5	12.2	244	1	US-07-869-933-32
22	123.5	12.2	244	1	US-08-201-879A-3
23	123.5	12.2	244	4	US-09-103-663-32
24	121.5	12.0	192	4	US-09-149-476-477
25	85	8.4	327	4	US-08-748-506-24
26	83.5	8.2	235	4	US-09-247-155-126
27	83	8.2	297	4	US-09-134-001C-5605

28	80.5	7.9	311	3	US-08-605-284B-23	Sequence 23, Appl
29	80.5	7.9	521	4	US-08-669-656A-4	Sequence 4, Appl
30	80.5	7.9	1856	4	US-08-843-417-2	Sequence 2, Appl
31	80.5	7.9	1957	4	US-08-669-656A-2	Sequence 2, Appl
32	80.5	7.9	1957	4	US-08-669-656A-8	Sequence 8, Appl
33	80.5	7.9	2132	4	US-08-669-656A-6	Sequence 6, Appl
34	79.5	7.8	192	4	US-09-026-017-2	Sequence 2, Appl
35	79.5	7.8	192	4	US-09-631-547-2	Sequence 2, Appl
36	79.5	7.8	241	3	US-08-808-148-1	Sequence 1, Appl
37	79.5	7.8	241	4	US-09-020-956-114	Sequence 114, App
38	79.5	7.8	241	4	US-09-030-607-114	Sequence 114, App
39	79.5	7.8	241	4	US-09-605-785-114	Sequence 114, App
40	79.5	7.8	241	4	US-09-439-313-114	Sequence 114, App
41	79.5	7.8	241	4	US-09-352-616A-114	Sequence 114, App
42	79.5	7.8	241	4	US-09-232-148A-114	Sequence 114, App
43	79.5	7.8	423	4	US-09-134-001C-5536	Sequence 5536, Ap
44	79	7.8	1872	6	5386025-6	Patent No. 5386025
45	79	7.8	1873	1	US-08-435-675B-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-08-916-902A-1
Sequence 1, Application US/08916902A
Patent No. 5871930
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HIGH AFFINITY IMMUNOGLOBULIN E
TITLE OF INVENTION: RECEPTOR-LIKE PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/916,902A
FILING DATE: Herewith
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0371 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 239 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAINOT04
CLONE: 927955
US-08-916-902A-1
Query Match 20.3%; Score 206.5; DB 2; Length 239;


```
US-08-707-340-4  
; Sequence 4, Application US/08707340  
; Patent No. 5705615  
; GENERAL INFORMATION:  
APPLICANT: Lim, Bing  
APPLICANT: Adra, Chaker N.  
TITLE OF INVENTION: RECOMBINANT HTM4 GENE, PROTEIN AND ASSAYS  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Millitia Drive  
CITY: Lexington  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02173  
  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/707,340  
FILING DATE: 03-SEP-1996  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/318,492  
FILING DATE: 06-OCT-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/675,648  
FILING DATE: 03-JUL-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Brook, David E.  
REGISTRATION NUMBER: 22,592  
REFERENCE/DOCKET NUMBER: BIH94-0342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 861-6240  
TELEFAX: (617) 861-9540  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 214 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-707-340-4
```

Query Match 18.2%; Score 184.5; DB 1; Length 214;
Best Local Similarity 28.2%; Pred. No. 5.2e-13;
Matches 59; Conservative 39; Mismatches 86; Indels 25; Gaps 8.

OY 1 MDSSIAH-SPEFLVPPPTTASVESTELSATFTSQSPLOKLFARKMKILTIQIIFGI 59
Db 1 LGSAAHHTPGSENGPELNTSVIH-----PINGSPTYQR--AKLVLCATITLNNA 60
OY - :|::||| |
Y- 60 MTFSFGVIFFLTLPKY---PRPFIFLSGYPMGVSYLEINSAGFLAVKKRTETTLII 115
Db 61 MILAGV-FLSLGYPHFQKHFEFTFYTGPIMGAVFCSSGLTSVAIGIKPTRTWIQ 119
OY - :|::||| |
Y- 116 LSRIIMNLISALRAIRAGIILLTFEGFLDONYI--CGYSRONSQ----CAAVTYLFIQLIT 169
Db 120 NSFGNINIASATIALGTAFSLINIVNIOQSRLSC---HSSESPLCLCNMGSIISGMWSL 176
OY - :|::||| |
Y- 170 LMFTSIELFILPSRLLSGSHEDDCQG 198
Db 177 LIILTELCVISTIANMCNNCCNSRE 205

RESULT 5
US-08-916-902A-3
; Sequence 3, Application US/08916902A
; Patent No. 5871930
; GENERAL INFORMATION:

[illegible]


```

      FILING DATE: 19920416
      CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
      NAME: BENT, Stephen A.
      REGISTRATION NUMBER: 29,768
      REFERENCE/DOCKET NUMBER: 40399/154 NTHD
TELECOMMUNICATION INFORMATION:
      TELEPHONE: (703)836-9300
      TELEFAX: (703)683-4109
      TELEX: 899149
      INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
      LENGTH: 243 amino acids
      TYPE: AMINO ACID
      TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
      ORGANISM: rat
      STRAIN: PCRI beta subunit
US-07-869-933-33

Query Match          15.7%; Score 159; DB 1; Length 243;
Best Local Similarity 28.0%; Pred. No. 4.5e-10;
Matches 45; Conservatly 29; Mismatches 71; Indels 16; Gaps 3.

QY       36 QSPLOKLPARKKIIGTIOILEGIMTFSGVFELFTLLKP--YRPFPFIETLSGYPFMGSV 93
|         | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db       47 QQTWOSFLKTELRFGLVGVQLICLCPFGIVWCSTLTOTSDPDEVLLIRAGYFPWGAV 106
|         | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY       94 LFINSAPLVAKRTTETTLILSRIMNLSALRAIAGIIITLTFGIIDONYIGCVSHON 153
|         | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

```

[illegible]

```

: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: GenBank
: CLONE: 204117
: US-08-916-902A-4

Query Match 15.7%; Score 159; DB 2; Length 243;
Best Local Similarity 28.0%; Pred. No. 4.5e-10;
Matches 45; Conservative 29; Mismatches 71; Indels 16; Gaps 3.

QY 36 QSPLOKLEARKKILGTQILFGIMTFSGVYFLETLKP--YRPFPIFLSGYFPMGSV 93
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 47 QQTMSFLKKELEFLGVYQVLGLCLCFGVYVCTLTQTSDEDDVLLIRAGYFPGAV 106
      || || | | | | | | | | | | | | | | | | | | | | | | | |
QY 94 LPIINGAFLLVAKRRTETLLISIMNLSALRAIAGIILLTFEGFILDQNYICYSQON 153
      || || | | | | | | | | | | | | | | | | | | | | | | | |
Db 107 LFIIVSGFISMSERKNTLYVRGSIKANIVSSIAAGLGIALLILMLSNMSAYM----- 159
      || || | | | | | | | | | | | | | | | | | | | | | | | |
QY 154 SOCKAVT-----VLFGLITLMTFETIIEFLISPPSIL 187
      : | | : | | : | | : | | : | | : | | : | | : | | : | | :
Db 160 NYCKDITEDGCEVTSTFTELVLMLFLTLAFCSAVLLITI 200
      : | | : | | : | | : | | : | | : | | : | | : | | : | | :

RESULT 12
US-09-213-389-4
: Sequence 4, Application US/09213389
: Patent No. 5977072
: GENERAL INFORMATION:
: APPLICANT: Bandman, Olga
: APPLICANT: Lal, Preeti
: APPLICANT: Corley, Neil C.
: TITLE OF INVENTION: HIGH AFFINITY IMMUNOGLOBULIN E
: TITLE OF INVENTION: RECEPTOR-LIKE PROTEIN
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: Incyte Pharmaceuticals, Inc.
: STREET: 3174 Porter Drive
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/213,389
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/916,902
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.
: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PF-0371 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-855-0555
: TELEFAX: 415-845-4166
: TELEX:
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 243 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: GenBank
: CLONE: 204117
: US-09-213-389-4

```

```

: LENGTH: 243
: TYPE: PROT
: ORGANISM: Rattus sp.
US-09-103-663-33

Query Match          15.7%; Score 159; DB 4; Length 243;
Best Local Similarity 28.0%; Pred. No. 4.5e-10;
Matches 45; Conservative 29; Mismatches 71; Indels 16; Gaps 3;

QY 36 QSPLOKLEFARKKILGTIOILFGIMTFSGVYFLFTLKP--YPRPPIFLSGYPPMGSV 93
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 47 QQTWQSFLEKKELEFGVYQVLGLCLCFGYVCCSTLTQSPDEDEVLLIYRAGYPPMGAV 106
      ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 94 LEINSGAFLLAVKRRTEETLIIISRNMLSLRAIAGIILLTFGFIIDONYICGYSQON 153
      ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 107 LEVLGSLFISMERKNTLYLVKSLGAINIVSSIAAGLAIALLIIMLNNSAVM----- 159
      ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 154 SOCKAVT-----VLPGLITLTSTSIIELEFISLPSIL 187
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 160 NYCKDITFEDGCFVYSFTTELVLMLFLTLIAFCSAVLLII 200
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 15
US-07-869-933-23
: Sequence 23, Application US/07869933
: Patent No. 5770396
: GENERAL INFORMATION:
: APPLICANT: KINET, Jean-Pierre
: TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF
: TITLE OF INVENTION: THE HUMAN B SUBUNIT OF THE HIGH AFFINITY RECEPTOR FOR
: NUMBER OF INVENTIONS: IMMUNOGLOBULIN
: NUMBER OF SEQUENCES: 34
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Foley & Lardner
: STREET: 1800 Diagonal Road, Suite 500
: CITY: Alexandria
: STATE: VA
: COUNTRY: USA
: ZIP: 22313-0299
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/869,933
: FILING DATE: 19920416
: CLASSIFICATION: 514
: ATTORNEY/AGENT INFORMATION:
: NAME: BENT, Stephen A.
: REGISTRATION NUMBER: 29,768
: REFERENCE/DOCKET NUMBER: 40399/154 NIHHD
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703)836-9300
: TELEFAX: (703)683-4109
: TELEX: 899149
: INFORMATION FOR SEQ ID NO: 23:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 246 amino acids
: TYPE: AMINO ACID
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-07-869-933-23

Query Match          15.7%; Score 159; DB 1; Length 246;
Best Local Similarity 28.0%; Pred. No. 4.6e-10;
Matches 45; Conservative 29; Mismatches 71; Indels 16; Gaps 3;

QY 36 QSPLOKLEFARKKILGTIOILFGIMTFSGVYFLFTLKP--YPRPPIFLSGYPPMGSV 93
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 50 QQTWQSFLEKKELEFGVYQVLGLCLCFGYVCCSTLTQSPDEDEVLLIYRAGYPPMGAV 109
      ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 94 LEINSGAFLLAVKRRTEETLIIISRNMLSLRAIAGIILLTFGFIIDONYICGYSQON 153

```

Db 110 LFLVLSGFLSISMRKNTLYLVGSLGANTVSSIAAGLAIILNLNNSAYM----- 162
OY 154 SOCKAVT-----VFLGILITLMTFSITELFISLPSIL 187
Db 163 NYCKDITEDDDGCFVTSFITELVIMLFLITIAFCNAVLLII 203

Search completed: February 24, 2003, 13:05:55
Job time : 23.8736 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2003, 13:04:41 : Search time 10.8649 Seconds
(without alignments)
569.068 Million cell updates/sec

Title: US-09-735-712-2

Perfect score: 1015
Sequence: 1 MDSTAHSPVFLVFPPEPTA.....ISLPSILGCHSDCCDCEQC 199

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 156504 seqs, 31069816 residues

Total number of hits satisfying chosen parameters: 156504

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEM_PUB pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCr_NEM_PUB pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEM_PUB pep:*
4: /cgn2_6/ptodata/1/pubpaa/US07_NEM_PUB pep:*
5: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB pep:*
6: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB pep:*
7: /cgn2_6/ptodata/1/pubpaa/PCrUS_PUBCOMB pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEM_PUB pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEM_PUB pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEM_PUB pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1015	100.0	199	10	US-09-735-712-2
2	1004	98.9	200	9	US-09-924-340-34
3	1004	98.9	200	9	US-09-992-600A-34
4	1004	98.9	200	10	US-09-821-821-2
5	1004	98.9	200	10	US-09-731-872-258
6	923	90.9	201	9	US-09-949-842-22
7	726.5	71.6	149	10	US-09-735-712-8
8	583	57.4	138	10	US-09-735-712-6
9	416	41.0	81	10	US-09-735-712-4
10	263.5	26.0	67	10	US-09-731-872-262
11	184.5	18.2	257	10	US-09-925-302-757
12	176.5	17.4	267	9	US-09-981-353-82
13	174.5	17.2	299	10	US-09-739-254-73
14	174.5	17.2	299	10	US-09-904-615-73
15	162.5	16.0	240	10	US-09-821-821-4
16	161.5	15.9	307	10	US-09-739-254-142
17	161.5	15.9	307	10	US-09-904-615-142
18	147.5	14.5	250	9	US-09-736-457-1677
19	147.5	14.5	250	9	US-09-902-941-1677

20	147.5	14.5	250	9	US-09-849-626-1677	Sequence 1677, Ap
21	147.5	14.5	302	10	US-09-925-2297-764	Sequence 764, App
22	142.5	14.0	167	10	US-09-765-205-22	Sequence 22, Appl
23	139.5	13.7	250	9	US-09-902-941-1874	Sequence 1874, Ap
24	139.5	13.7	250	9	US-09-849-626-1874	Sequence 1874, Ap
25	139.5	13.7	286	9	US-09-902-941-1878	Sequence 1878, Ap
26	139.5	13.7	286	9	US-09-849-626-1878	Sequence 1878, Ap
27	139.5	13.7	384	9	US-09-902-941-1876	Sequence 1876, Ap
28	139.5	13.7	384	9	US-09-849-626-1876	Sequence 1876, Ap
29	133	13.1	204	10	US-09-739-254-117	Sequence 117, App
30	133	13.1	204	10	US-09-904-615-117	Sequence 117, App
31	128.5	12.7	155	9	US-09-796-692-1642	Sequence 1642, Ap
32	115.5	11.4	149	9	US-09-796-692-1077	Sequence 1077, Ap
33	99.5	9.8	94	9	US-09-796-692-877	Sequence 877, App
34	98.5	9.7	125	9	US-09-796-692-1088	Sequence 1088, Ap
35	98.5	9.7	125	9	US-09-796-692-1562	Sequence 1562, Ap
36	98	9.7	311	10	US-09-925-297-788	Sequence 788, App
37	87	8.6	210	9	US-09-862-540-54	Sequence 54, Appl
38	85	8.4	302	10	US-09-771-730-62	Sequence 62, App
39	83.5	8.2	236	9	US-09-981-876-202	Sequence 202, App
40	83.5	8.2	236	9	US-09-148-545-202	Sequence 202, App
41	83	8.2	331	9	US-09-965-529-33	Sequence 33, Appl
42	82.5	8.1	228	10	US-09-815-242-11542	Sequence 11542, A
43	80	7.9	240	9	US-09-908-193-43	Sequence 43, Appl
44	79.5	7.8	192	10	US-09-815-242-12330	Sequence 12330, A
45	79.5	7.8	192	10	US-09-815-242-12791	Sequence 12791, A

ALIGNMENTS

RESULT 1
US-09-735-712-2
; Sequence 2, Application US/09735712
; Patent No. US20020045743A1
; GENERAL INFORMATION:
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20020045743A1 Human Membrane Proteins and
; FILE REFERENCE: Polynucleotides Encoding the Same
; CURRENT APPLICATION NUMBER: US/09/735,712
; CURRENT FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 60/171,567
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 199
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-735-712-2

Query Match	100.0%;	Score 1015;	DB 10;	Length 199;
Best Local Similarity	100.0%;	Pred. No. 1.7e-94;		
Matches 199;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MDSTAHSPVFLVFPPEPTA	SESTETLSATFTSPLOKLPARKMKTLGTQILFGIM	60
DB	1	MDSTAHSPVFLVFPPEPTA	SESTETLSATFTSPLOKLPARKMKTLGTQILFGIM	60
QY	61	TFSFGVIFLTLLKPYPRPF	FLSGYPWGSVLFNSGAFLLAVKRKTELLILSRIM	120
DB	61	TFSFGVIFLTLLKPYPRPF	FLSGYPWGSVLFNSGAFLLAVKRKTELLILSRIM	120
QY	121	NLSALRAIAGIILLTFGFI	LDONTCGYSHNSOCKATVFLGILITLMTFTSIELEFI	180
DB	121	NLSALRAIAGIILLTFGFI	LDONTCGYSHNSOCKATVFLGILITLMTFTSIELEFI	180
QY	181	SLPFSILGCHSDCCDCEQC	199	
DB	181	SLPFSILGCHSDCCDCEQC	199	

```
RESULT 2
US-09-924-340-34
; Sequence 34, Application US/09924340
; Publication No. US20030027248A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US2.REG
; CURRENT APPLICATION NUMBER: US/09/924,340
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: Jpatent
; SEQ ID NO 34
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-924-340-34

Query Match          98.9%; Score 1004; DB 9; Length 200;
Best Local Similarity 99.0%; Pred. No. 2.2e-93;
Matches 197; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDSSTAHSPVFLVPPETASVESTELSATFTSTQSPLOKLFARKMKILGTIQLFGIM 60
    |||||||
DB 1 MDSSTAHSPVFLVPPETASVESTELSATFTSTQSPLOKLFARKMKILGTIQLFGIM 60
    |||||||
QY 61 TFSFGVIFLFTLKPYPRPFIFLSGYPFGSVLFINSGAFLIAVKRRTETLLILSRIM 120
    |||||||
DB 61 TFSFGVIFLFTLKPYPRPFIFLSGYPFGSVLFINSGAFLIAVKRRTETLLILSRIM 120
    |||||||
QY 121 NLSALRAIAGIILTFEFLIDQNYICGYSHONSCKAVTVLFGILITLMTFSIIELEFI 180
    |||||||
DB 121 NLSALRAIAGIILTFEFLIDQNYICGYSHONSCKAVTVLFGILITLMTFSIIELEFI 180
    |||||||
QY 181 SLPFSIILGCHSEDCDCEQC 199
    |||||||
b 181 SLPFSIILGCHSEDCDCEQC 199
    |||||||

RESULT 3
US-09-992-600A-34
; Sequence 34, Application US/0992600A
; Publication No. US20030027161A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US4.DIV
; CURRENT APPLICATION NUMBER: US/09/992,600A
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/01715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 114
```

```
; SOFTWARE: Jpatent
; SEQ ID NO 34
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-992-600A-34

Query Match          98.9%; Score 1004; DB 9; Length 200;
Best Local Similarity 99.0%; Pred. No. 2.2e-93;
Matches 197; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDSSTAHSPVFLVPPETASVESTELSATFTSTQSPLOKLFARKMKILGTIQLFGIM 60
    |||||||
DB 1 MDSSTAHSPVFLVPPETASVESTELSATFTSTQSPLOKLFARKMKILGTIQLFGIM 60
    |||||||
QY 61 TFSFGVIFLFTLKPYPRPFIFLSGYPFGSVLFINSGAFLIAVKRRTETLLILSRIM 120
    |||||||
DB 61 TFSFGVIFLFTLKPYPRPFIFLSGYPFGSVLFINSGAFLIAVKRRTETLLILSRIM 120
    |||||||
QY 121 NLSALRAIAGIILTFEFLIDQNYICGYSHONSCKAVTVLFGILITLMTFSIIELEFI 180
    |||||||
DB 121 NLSALRAIAGIILTFEFLIDQNYICGYSHONSCKAVTVLFGILITLMTFSIIELEFI 180
    |||||||
QY 181 SLPFSIILGCHSEDCDCEQC 199
    |||||||
DB 181 SLPFSIILGCHSEDCDCEQC 199
    |||||||

RESULT 4
US-09-821-821-2
; Sequence 2, Application US/09821821
; Patent No. US20020064823A1
; GENERAL INFORMATION:
; APPLICANT: Welcher, Andrew A.
; APPLICANT: Calzone, Frank J.
; TITLE OF INVENTION: CD20/1gE-Receptor Like Molecules and Uses thereof
; FILE REFERENCE: 01017/36938A
; CURRENT APPLICATION NUMBER: US/09/821,821
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: US 09/723,258
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: US 60/193,728
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-821-821-2

Query Match          98.9%; Score 1004; DB 10; Length 200;
Best Local Similarity 99.0%; Pred. No. 2.2e-93;
Matches 197; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDSSTAHSPVFLVPPETASVESTELSATFTSTQSPLOKLFARKMKILGTIQLFGIM 60
    |||||||
DB 1 MDSSTAHSPVFLVPPETASVESTELSATFTSTQSPLOKLFARKMKILGTIQLFGIM 60
    |||||||
QY 61 TFSFGVIFLFTLKPYPRPFIFLSGYPFGSVLFINSGAFLIAVKRRTETLLILSRIM 120
    |||||||
DB 61 TFSFGVIFLFTLKPYPRPFIFLSGYPFGSVLFINSGAFLIAVKRRTETLLILSRIM 120
    |||||||
QY 121 NLSALRAIAGIILTFEFLIDQNYICGYSHONSCKAVTVLFGILITLMTFSIIELEFI 180
    |||||||
DB 121 NLSALRAIAGIILTFEFLIDQNYICGYSHONSCKAVTVLFGILITLMTFSIIELEFI 180
    |||||||
QY 181 SLPFSIILGCHSEDCDCEQC 199
    |||||||
DB 181 SLPFSIILGCHSEDCDCEQC 199
    |||||||

RESULT 5
```



```
US-09-731-872-258
; Sequence 258, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78 US3 REG
; CURRENT APPLICATION NUMBER: US/09/731,872
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 258
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -20...-1
US-09-731-872-258
```

```
Query Match          98.9%; Score 1004; DB 10; Length 200;
Best Local Similarity 99.0%; Pred. No. 2.2e-93;
Matches 197; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 MDSTASHPVFLVFPPEITASEYESTELSATFTSTOSPLOKLPARKKILGTIQLFGIM 60
    |||||||
DB 1 MDSTASHPVFLVFPPEITASEYESTELSATFTSTOSPLOKLPARKKILGTIQLFGIM 60
    |||||||
QY 61 TFSFGVLEFTLLKPYRPFIFLSGYPFWGVLFINSGAFLAVKRTETLILSRIM 120
    |||||||
DB 61 TFSFGVLEFTLLKPYRPFIFLSGYPFWGVLFINSGAFLAVKRTETLILSRIM 120
    |||||||
QY 121 NLSALRAIGIILLTFGLFDONICGYSHONSOKAVTVLFGILITLMTFSIIEFLT 180
    |||||||
DB 121 NLSALRAIGIILLTFGLFDONICGYSHONSOKAVTVLFGILITLMTFSIIEFLT 180
    |||||||
QY 181 SLPFSILGCHSEDCDEQC 199
    |||||||
DB 181 SLPFSILGCHSEDCDEQC 199
    |||||||
```

```
RESULT 6
US-09-949-842-22
; Sequence 22, Application US/09949842
; Patent No. US20020164692A1
; GENERAL INFORMATION:
; APPLICANT: N1 et al.
; TITLE OF INVENTION: PT047P1
; FILE REFERENCE: Immune System-Related Polynucleotides, Polypeptides, and Antibodies
; CURRENT APPLICATION NUMBER: US/09/949,842
; CURRENT FILING DATE: 2001-09-02
; PRIOR APPLICATION NUMBER: PCT/US01/07260
; PRIOR FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: 60/224,367
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/187,873
; PRIOR FILING DATE: 2000-03-08
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-949-842-22
```

```
Query Match          90.9%; Score 923; DB 9; Length 201;
Best Local Similarity 98.4%; Pred. No. 2.8e-85;
```

```
Matches 185; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 MDSTASHPVFLVFPPEITASEYESTELSATFTSTOSPLOKLPARKKILGTIQLFGIM 60
    |||||||
DB 1 MDSTASHPVFLVFPPEITASEYESTELSATFTSTOSPLOKLPARKKILGTIQLFGIM 60
    |||||||
QY 61 TFSFGVLEFTLLKPYRPFIFLSGYPFWGVLFINSGAFLAVKRTETLILSRIM 120
    |||||||
DB 61 TFSFGVLEFTLLKPYRPFIFLSGYPFWGVLFINSGAFLAVKRTETLILSRIM 120
    |||||||
QY 121 NLSALRAIGIILLTFGLFDONICGYSHONSOKAVTVLFGILITLMTFSIIEFLT 180
    |||||||
DB 121 NLSALRAIGIILLTFGLFDONICGYSHONSOKAVTVLFGILITLMTFSIIEFLT 180
    |||||||
QY 181 SLPFSILG 188
    |||||||
DB 181 SLPFSIWG 188
    |||||||
```

```
RESULT 7
US-09-735-712-8
; Sequence 8, Application US/09735712
; Patent No. US20020045743A1
; GENERAL INFORMATION:
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Walke, D. Wade
; TITLE OF INVENTION: NO. US20020045743A1 Human Membrane Proteins and
; FILE REFERENCE: Polynucleotides Encoding the Same
; CURRENT APPLICATION NUMBER: US/09/735,712
; CURRENT FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 60/171,567
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-735-712-8
```

```
Query Match          71.6%; Score 726.5; DB 10; Length 149;
```

```
Best Local Similarity 74.4%; Pred. No. 9.4e-66;
Matches 148; Conservative 0; Mismatches 0; Indels 51; Gaps 1;
```

```
QY 1 MDSTASHPVFLVFPPEITASEYESTELSATFTSTOSPLOKLPARKKILGTIQLFGIM 60
    |||||||
DB 1 MDSTASHPVFLVFPPEITASEYESTELSATFTSTOSPLOKLPARKKILGTIQLFGIM 60
    |||||||
QY 61 TFSFGVLEFTLLKPYRPFIFLSGYPFWGVLFINSGAFLAVKRTETLILSRIM 120
    |||||||
DB 61 TFSFGVLEFTLLKPYRPFIFLSGYPFWGVLFINSGAFLAVKRTETLILSRIM 120
    |||||||
QY 121 NLSALRAIGIILLTFGLFDONICGYSHONSOKAVTVLFGILITLMTFSIIEFLT 180
    |||||||
DB 113 -----LGLITLMTFSIIEFLT 129
    |||||||
```

```
RESULT 8
US-09-735-712-6
; Sequence 6, Application US/09735712
; Patent No. US20020045743A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20020045743A1 Human Membrane Proteins and
; FILE REFERENCE: Polynucleotides Encoding the Same
; CURRENT APPLICATION NUMBER: US/09/735,712
```

```
; CURRENT FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 60/171,567
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 138
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-735-712-6

Query Match
Best Local Similarity 57.4%; Score 583; DB 10; Length 138;
Matches 117; Conservative 5; Mismatches 0; Indels 6; Gaps 1;

QY 1 MDSSTASHPVFLVFPPEITASEYESTELSATFTSTQSPLOKLFARKMKILGTIQLIFGIM 60
DB 1 MDSSTASHPVFLVFPPEITASEYESTELSATFTSTQSPLOKLFARKMKILGTIQLIFGIM 60
QY 61 TFSFGVIFLFTLLKPYRPFPIFISGYPFGSVLFINSAGFLAVARKTETELI----- 114
DB 61 TFSFGVIFLFTLLKPYRPFPIFISGYPFGSVLFINSAGFLAVARKTETELI----- 114
QY 115 ILSRIMNL 122
DB 121 ILTKIVSV 128

RESULT 9
US-09-735-712-4
; Sequence 4, Application US/09735712
; Patent No. US20020045743A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20020045743A1e1 Human Membrane Proteins and
; TITLE OF INVENTION: Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0109-USA
; CURRENT APPLICATION NUMBER: US/09/735,712
; CURRENT FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 60/171,567
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 81
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-735-712-4

Query Match
Best Local Similarity 41.0%; Score 416; DB 10; Length 81;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 MNLSLALRAIAGIILTLTFCFILDONTICGSHONSCKAVTVLFGILITLMTFSIIELF 179
DB 1 MNLSLALRAIAGIILTLTFCFILDONTICGSHONSCKAVTVLFGILITLMTFSIIELF 60
QY 180 ISLPSILICHSBDCCEOC 199
DB 61 ISLPSILICHSBDCCEOC 80

RESULT 10
US-09-731-872-262
; Sequence 262, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Malne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.053.REG
```

```
; CURRENT APPLICATION NUMBER: US/09/731,872
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 262
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -20...-1
US-09-731-872-262

Query Match
Best Local Similarity 26.0%; Score 263.5; DB 10; Length 67;
Matches 56; Conservative 3; Mismatches 7; Indels 9; Gaps 1;

QY 1 MDSSTASHPVFLVFPPEITASEYESTELSATFTSTQSPLOKLFARKMKILGTIQLIFGIM 60
DB 1 MDSSTASHPVFLVFPPEITASEYESTELSATFTSTQSPLOKLFARKMKILGDIH----- 54
QY 61 TFSFGVIFLFTLLKPY 75
DB 55 ---SGALFCSLIEP 66

RESULT 11
US-09-925-302-757
; Sequence 757, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 757
; LENGTH: 257
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (210)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-302-757

Query Match
Best Local Similarity 18.2%; Score 184.5; DB 10; Length 257;
Matches 55; Conservative 40; Mismatches 90; Indels 23; Gaps 6;

QY 1 MDSSTASHPVFLVFPPEITASEYESTELSATFTSTQSPLOKLFARKMKILGTIQLIFGIM 60
DB 26 MTSQVPVNETIIVLPSNVI--NFSQAEKPEPTNOCODSLKHLHAIRKIVGTIQLICGWM 83
QY 61 TFSFGVIFLFTLLKPYRPFPIFL-SCYPPFGSVLFINSAGFLAVARKTETELIILSR 118
DB 84 VLSGLITLASASFSPFTQVSTLNSAVPFLGPFPIISGLSLATEKRLKLVHSSL 143
QY 119 INMLSLALRAIAGIILTLTFCFILDONTICGSHONSCKAVTVLFGILITLMTFSIIELF 159
DB 144 VCSIIISALVALGFIILSKQATLPASLOCCLDKNMIPTRSYVYFHDLSYTTDCYTA 203
QY 160 TVLFLGILITLMTFSIIELFISLPSIL 187
```


; PRIOR APPLICATION NUMBER: US 09/723,258
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: US 60/193,728
 ; PRIOR FILING DATE: 2000-03-30
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 4
 ; LENGTH: 240
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-821-821-4

Query Match 16.0%; Score 162.5; DB 10; Length 240;
 Best Local Similarity 25.6%; Pred. No. 5e-09;
 Matches 50; Conservative 40; Mismatches 76; Indels 29; Gaps 7;

QY 15 PPEITASEYESTELSAFTSTQSPLOKLFARKMKILGTIQLFGIMTFSGVIFLFTLK 74
 Db 16 PKGITIPQREK---PGHMYQNEEDYLQNGLPETETVLGTVOILCCLLSSGAILVFA--- 69
 QY 75 PYPRF--PFI---FLSGYFPGSVLFINSGAFLIAVRKKTETIILSRIMNLISALRAI 129
 Db 70 PYSHENPAISTYLMISGYPFLGALCFGITGSLTISGKQSTKPPDLSSLTSMNAVSVYTAG 129
 QY 130 AGIILTFGFTL-----DQNYICG--YS-----HONSQKAVTVLFGILITM 171
 Db 130 AGFLADSMVALATASQHGSGEMDYLSLPYSEYYPPIYEIKDCLLTSVSLGVLVVM 189
 QY 172 TFSIIELEFISLPFSI 186
 Db 190 IFVLELLLAAYSSV 204

Search completed: February 24, 2003, 13:11:52
 Job time : 10.8649 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2003, 13:01:21 : Search time 16.5833 Seconds

(without alignments)
1153.613 Million cell updates/sec

Title: US-09-735-712-2

Perfect score: 1015

Sequence: 1 MDSTSHSPVFLVFPPEITR.....ISLPSILGCHSEDCDEQC 199

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR_73:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1004	98.9	200	2 JC7585	testis expressed t
2	184.5	18.2	214	2 I59258	IgE receptor beta
3	112	16.9	291	2 A30558	B-cell surface ant
4	168	16.6	297	1 A30586	B-cell surface ant
5	159	15.7	243	2 A31231	high-affinity IgE
6	158.5	15.6	235	2 B34342	IgE Fc receptor be
7	123.5	12.2	244	2 A43806	IgE Fc receptor be
8	103	10.1	307	2 C86844	IgE receptor beta
9	100	9.9	152	2 JH0751	IgE receptor beta
10	96	9.5	387	2 E81306	probable transmem
11	96	9.5	455	2 T16070	hypothetical prote
12	95.5	9.4	537	2 T07892	probable inorganic
13	95	9.4	408	2 C70379	hypothetical prote
14	94.5	9.3	457	2 T37205	hypothetical prote
15	94	9.3	439	2 E55155	hypothetical prote
16	94	9.3	439	2 H98021	hypothetical prote
17	94	9.3	770	2 G72589	hypothetical prote
18	93.5	9.2	542	2 T01124	hypothetical prote
19	92.5	9.1	230	2 S50330	probable phosphate
20	92.5	9.1	533	2 T07894	H+-transporting tw
21	92	9.1	324	2 D30020	NADH2 dehydrogenas
22	91.5	9.0	666	2 S59109	NADH2 dehydrogenas
23	90.5	8.9	694	2 E69143	hypothetical prote
24	90	8.9	165	2 F69208	hypothetical prote
25	90	8.9	353	2 T42971	hypothetical prote
26	89	8.8	288	2 A05235	hypothetical prote
27	89	8.8	485	2 E98261	L-asparagine perme
28	89	8.8	485	2 AD3023	L-asparagine perme
29	88.5	8.7	538	2 T05714	probable inorganic

30	88.5	8.7	538	2 T07164	probable inorganic
31	88.5	8.7	540	2 T07604	phosphate transpor
32	88	8.7	141	2 H89831	hypothetical prote
33	87.5	8.6	170	2 C90079	conserved hypothet
34	87.5	8.6	381	2 F59101	hypothetical prote
35	87.5	8.6	398	2 AF0049	probable membrane
36	87	8.6	147	2 H89963	conserved hypothet
37	87	8.6	312	2 S01191	NADH2 dehydrogenas
38	87	8.6	382	2 C36966	O-antigen polymera
39	87	8.6	481	2 T23131	hypothetical prote
40	87	8.6	555	2 F72555	probable molypdenu
41	86	8.5	212	2 A99370	hypothetical prote
42	86	8.5	880	2 D69427	conserved hypothet
43	85.5	8.4	202	2 D90499	conserved hypothet
44	85.5	8.4	319	2 H69160	hypothetical prote
45	85.5	8.4	458	1 YTB5RT	tetracycline resis

ALIGNMENTS

```
RESULT 1
JC7585
testis expressed transmembrane-4 protein, TETM4 - human
C:Species: Homo sapiens (man)
C:Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 30-Jun-2001
C:Accession: JC7585
R:Hulet, M.D.; Pagier, E.; Hornby, J.R.; Hogarth, P.M.; Eyre, H.J.; Baker, E.; Crawford,
Biochem. Biophys. Res. Commun. 280, 374-379, 2001
A:Title: Isolation, tissue distribution, and chromosomal localization of a novel test
A:Reference number: JC7585; MIMD:21092614; PMID:11162526
A:Accession: JC7585
A:Molecule type: mRNA
A:Residues: 1-200 <HUL>
A:Cross-references: GB:AF311127
C:Comment: This protein, a four-transmembrane protein, associates with receptor comp
C:Genetics:
A:Gene: tetm4
A:Map position: 11q12
C:Keywords: signal transduction
F:1-48/Domain: cytoplasmic #status predicted <CYT1>
F:49-70/Domain: transmembrane #status predicted <TM1>
F:71-84/Domain: extracellular #status predicted <EX1>
F:85-105/Domain: transmembrane #status predicted <TM2>
F:106-118/Domain: intracellular #status predicted <INT>
F:119-138/Domain: transmembrane #status predicted <TM3>
F:139-160/Domain: extracellular #status predicted <EX2>
F:161-182/Domain: transmembrane #status predicted <TM4>
F:183-200/Domain: cytoplasmic #status predicted <CYT2>

Query Match          98.9%; Score 1004; DB 2; Length 200;
Best Local Similarity 99.0%; Pred. No. 1.7e-87;
Matches 197; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDSTSHSPVFLVFPPEITRSEYESTLSTPTSTOSPILOKLPARKKILGTIOILFGIM 60
DB 1 MDSTSHSPVFLVFPPEITRSEYESTLSTPTSTOSPILOKLPARKKILGTIOILFGIM 60
QY 61 TFSFGVIFLTLLKPYPRPFIFLISGYPFGSVLFINSGAFLAVRKTETTLISRLM 120
DB 61 TFSFGVIFLTLLKPYPRPFIFLISGYPFGSVLFINSGAFLAVRKTETTLISRLM 120
QY 121 NLSALRAIAGIILTFGFIIDONYICGYSHQSGKRAVTVLFGILITLMTFSIELFT 180
DB 121 NLSALRAIAGIILTFGFIIDONYICGYSHQSGKRAVTVLFGILITLMTFSIELFT 180
QY 181 SLFPSILGCHSEDCDEQC 199
DB 181 SLFPSILGCHSEDCDEQC 199

RESULT 2
I59258
```

DY 91 GSVLFINSAGFLIAVRKRTETLLIIISR-IMNLSALPAIAGITL-----LTFFGFIID 142
|::|||::|||::|||::|||
Db 83 GGIMYIIISGLLAATAAEKTISKRSIVAKVIMSSLSFPAISGITLIMDILNMTHSHFLK 142

Dd 36 SLVGPDSFFMRESKTLGAVQIMNGLFHLALGL---LMIPIAGIYAPICVTYVWYPLMG 91

DQ 34 STQSPLOKLFARKKKILGTIQLFGIMTFSCVIFLETLTKPPRF-PRFILSGYPMWG 92
| | | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 36 SLVGPTQSFFMRSEKTLGAVQIMNGLFLIALGL---LMIPAGIYAPICVTWVPPLMG 91

Db 92 IMYIISSGLAATERNSKCLVORKMIMNSLSLEPAALSGMLSIMDILINKIKISFLKMES 151

QY 137 FGFILDON-YICGY-----SHONS-----QCKAVTVLFGILITLTFESIT-ELFIS 181

Db 152 LMFRAHPTNYININCEPANSENKNSSTQVCYSIGSLFGILISVMILIFAPFQELAVYA 209

RESULT 5

high-affinity IGE receptor beta chain - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 05-Nov-1999
C:Accession: A31231
R:Kinetic, J.P.; Blank, U.; Ra, C.; White, K.; Metzger, H.; Kochan, J.
Proc. Natl. Acad. Sci. U.S.A. 85, 6483-6487, 1988
A:Title: Isolation and characterization of cDNAs coding for the beta-subunit of the high-affinity IgE receptor
A:Reference number: A31231; MUID:88320465; PMID:2970642
A:Accession: A31231
A:Molecule type: mRNA
A:Residues: 1-243 <K>IN
A:Cross-references: GB:W22923; GB:J03845; NID:g204116; PIDD:AAAA1149.1; PID:g204117
C:Keywords: Immunoglobulin receptor; Transmembrane protein

Query Match 15.7%; Score 159; DB 2; Length 243;
Best Local Similarity 28.0%; Pred. No. 1.2e-07;
Matches 45; Conservative 29; Mismatches 71; Indels 16; Gaps 3;

QY 36 QSPLOKLFARKMKILGTIQLFGIMTFSGVIFLFTLLKP--YPRPFILSGYPFWGSV 93

Db 47 QQTWQSLKKELEFLGVTQVLVGLICLQFGTVVCSLTQTSDFDEVELLLYRAGIYFWGAV 106

94 LFINGGAFLIAVKRTTETLLILSRIMNULSALRAIGILLTGFILDQNYICGYSHQ 153

Db 107 LEVLGFLSMERKNTLYLVKSLGANTVSSIAAGLGIALLNLNSNSAYM----- 159

OV 154 SOCKAVT-----VI.FIGII.TTIMTESTIEI.FISI.PESI. 187

```

160 NVCKRTMEDDCCEVETSEIEMI MI I EITTI AECGAVITI 300
      : | : | : : | | | | :

```

RESULT 6

IgE Fc receptor/beta chain - mouse
C:Species: Mus musculus (house mouse)
C:Date: 15-Jun-1990 #sequence_revision 15-Jun-1990 #text_change 05-Nov-1999
C:Accession: B34342
R:Ra, C.; Jouvin, M.H.E.; Kinet, J.P.
J. Biol. Chem. 264, 15323-15327, 1989
A:Title: Complete structure of the mouse mast cell receptor for IgE (Fc-epsilon-RI) and
A:Reference number: A34342; MUID:89359361; PMID:3527850

A; Cross-references: GB:J05019; NID:g193238; PIDN:AAA37601.1; PID:g309225
C; Keywords: immunoglobulin receptor; transmembrane protein

Query Match	15.6%	Score 158.5;	DB 2;	Length 235;
Best Local Similarity	30.58%	Pred. No. 1.3e-07;		
Matches 60; Conservative	29;	Mismatches 81;	Indels 27;	Gaps 9;

QY 1 MDSTAHSPVFLVFP-PEITASEYESTELSATTFSTQSPLO--KLFARK-MKILGIIQIL 56

```
Db      | | | : : | | : | | : | : | |
| : | | | : : | | : | | : | |
1 MDTEN-RSRADIALPNPOESSAPDIFILEASPAKAPPKOMWRFETKKLEFETGATOTT, 59
```

57 PCIMESEGVTEL ENOI I KDV - - - - DBEDEL EGVDEACGV EITNCGAF IAVUDRTMB 11

[illegible]

DB 60 VGLICICFGTIVCSYL---YSDFDDEEVLLTKLGIPFWGAVLFVLSGFLSIISERNITL 11

QY 112 TLITSRIMULSALRAIAGIITLTFEGFILDQNYICGYSHQNSQCAVT-----VFL 16

Db 117 YLVKSLGANIVSSIAAGTGIALIL-----NLTNFAFYMNN-CKNTEDEDCFEVASFT 16

```
QY      165 GILITMTESIIEFIS 181
          | : : | | | |
Db      170 TELVIMMLEFTIIFCS 186
```

RESULT 7

I:GE_Fc receptor beta chain - human
C:Species: Homo sapiens (man)
C:Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 08-Oct-1999
C:Accession: A42806; S21154
R:Kuester, H.; Zhang, L.; Brini, A.T.; MacGlashan, D.W.J.; Kinet, J.P.
J. BIOL. Chem. 267, 12782-12787, 1992
A:Title: The gene and cDNA for the human high affinity immunoglobulin E receptor beta
A:Reference number: A42806; MUID:92316966; PMID:1535625
A:Accession: A42806
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-244 <KUE>
A:Cross-references: GB:069796; NID:g337417; PIDD:AAA60269.1; PID:g337418
R:Maekawa, K.; Imagawa, N.; Tanaka, Y.; Harada, S.

A1:Title: Determination of the sequence coding for the beta subunit of the human high A1:Reference number: S21154; MID:92339505; PMID:1386024

A:Status: preliminary
A:Accession: S21154
A:Molecule type: mRNA

A:Residues: 1-244 <MAE>
A:Cross-references: GB:D10583; NID:g219881; PIDN:BAA01440.1; PID:g219881
A:Cross-references: GB:D10583; NID:g219881; PIDN:BAA01440.1; PID:g219881

C;Keywords: immunoglobulin receptor; transmembrane proteins
A;Introns: 19/2; 62/3; 107/3; 126/3; 179/3; 212/3
C;Genetics:

Query Match	12.28;	Score 123.5;	DB 2;	Length 244;
Best Local Similarity	25.68;	Pred. No. 0.00027;		

Matches 50; Conservative 33; Mismatches 93; Indels 19; Gaps 7;

QY 9 PVFLVFPPEITASEYESTELSATTFSTQSPLO---KLFARKMKILCTIQLFGIMTFSEFG

Db 21 PAFEVL--EISPQEVSSGRLLKS--ASSPPLHTWLTVLKKEQEFGLGVITQILTMAMICLCFG

66 VIFLFTLLKPYRPF--FLSGYPWGSVLFNSGAFLIAVKKRTTETLIILSRIMLL 123

```

: | : | | | | | : | | : | : |
77 TWCVSVIDISHTECHESSEKACYPENCLTEFSTSCMSTISEBNATYVVRGSGICANTA 136

```

[illegible]

```

QY 124 SAKRAIAGILLEHFGILBQNYICGYSHQNS-----QCKAWIVELSLITLMTF-SILE 177
      ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::

```

Db 137 SSIAAGTGITIL---IINLKSLAYIHHSQKFFETKCFMASFSTEIVVMLEFLIIG 1922

QY 178 LFISLPFSLGCHSE 192

hypothetical protein ysbC [imported] - Lactococcus lactis subsp. lactis (strain IL14)
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: C68844
R:Boitron, A.; Minkler, P.; Manger, S.; Jallion, O.; Malmgren, K.; Weissbach, J.; Eberhard
Genome Res. 11, 751-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis
#:Reference number: A86625; MUID:21235186; PMID:11337471

A;Accession: C86844
A;Status: preliminary

A: Molecule type: DNA
A: Residues: 1-307 <STO>

A/Cross-references: GB:AE005176; PID:q12724776; PIDN:AAK05853.1; GSPDB:GN00146
A/Environmental sources: strain T17A03

A/Experimental source: strain 1B1403
C/Genetics:

A;Gene: ysbC

Db 432 --IPPARLRSTCHGISAAAGKAGAVGAFGLVAAQSKDPTKDKGYPTGIGIKNSLIML 489
 QY 159 VTVEFLGILITLM 171
 Db 490 GVINEVGMCLTLL 502

RESULT 13

C70379
 hypothetical protein ag_917 - Aquifex aeolicus
 C/Species: Aquifex aeolicus
 C/Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
 C/Accession: C70379
 R/Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Oy

V.
 Nature 392, 353-358, 1998
 A/Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
 A/Reference number: A70300; MID:98196666; PMID:9537320
 A/Accession: C70379

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-408 <AQF>
 A/Cross-references: GB:AE000713; MID:g2983424; PIDN:AAC07022.1; PID:g2983440; GB:AE00065

A/Experimental source: strain VF5

C/Genetics:

A/Gene: ag_917

Query Match

Best Local Similarity 25.1%; Pred. No. 0.22;

Matches 51; Conservative 23; Mismatches 45; Indels 84; Gaps 10;

QY 36 QSPLOK-LFARKKILGTIILGIMFES---FGVLEFLTLKPYRPFIFLSGYPW 90
 Db 79 EDPLKATFVHIGTLTLVGLVSNVALLGLFLLI----- 119
 QY 91 GSVLFINSAGFLAVKRRKTEETILISIMNLSALRAI-----AGIILLT----- 136
 Db 120 -SLYXISS---LMLYK-----LIRIKNLNDAPRGFPFALGSFTAGIILATFVLNL 166
 QY 137 FGRF-LDQNT-----ICGYSHO-----NSQCKAVTVLF 163
 Db 167 FGFNLHKKYLFDFLHMSFMLEGWTATLVASVFSQVIEFVTPPYPKFISSYLEKTVETL 226
 QY 164 LGILITMTFSIIEFLISLPFSI 186
 Db 227 LVTKVFLPNESLIDVFISLIIFTI 249

RESULT 14

T37205
 hypothetical protein y57g7a.4 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C/Accession: T37205
 R/Ozersky, P.
 submitted to the EMBL Data Library, March 1999
 A/Description: The sequence of C. elegans cosmid y57g7a.
 A/Reference number: Z21634
 A/Accession: T37205

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-457 <OZE>

A/Cross-references: EMBL:AF077542; PIDN:AAC26295.1

C/Genetics:

A/Map position: II

A/Intons: 100/2; 193/2; 234/3; 272/2; 313/2; 399/1

A/Note: Y57G7A.4

Query Match

Best Local Similarity 25.0%; Pred. No. 0.28; Mismatches 48; Conservative 25; Mismatches 62; Indels 57; Gaps 8;

QY 14 FPEITASEYESTELSATFTSPLOKLPARKKILGTIILGIMTFSFGVLEFLTLL 73

Db 197 FPNVITSEYMTAKKAYSRADOTPLLVFI-----ILDGLSKITPAVI----- 240
 QY 74 KYPKPFPIFLSGYFPGSVLFINSAGFLAVKRRKTEETILISIMNLSALRAIAGII 133
 Db 241 -YPIFTF-----LLIQURLAAIARRKRTSTSM--GSRLSTKSD-QTFKVI 283
 QY 134 LITPFGIILDQNYICGYSHQNSQCKAVTVLFILITMTFSIE-----LEISLPFS 185
 Db 284 LMTVTFIISGPI-----GICYILEGTLPKRSVPFDINYLMDAFTIFVAINAS 332
 QY 186 I-----IGCHSE 192
 Db 333 VHFILCLGVHSQ 344

RESULT 15

E95155
 hypothetical protein SPI340 [imported] - Streptococcus pneumoniae (strain TIGR4)
 C/Species: Streptococcus pneumoniae
 C/Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
 C/Accession: E95155
 R/Tetelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; P
 on, J.D.; Umeyam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtap
 son, T.; Hickey, E.K.; Holt, I.E.
 Science 293, 498-506, 2001

A/Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morris

A/Title: Complete Genome Sequence of a Virulent Isolate of Streptococcus pneumoniae.

A/Reference number: A95000; MID:21357209; PMID:11463916

A/Accession: E95155

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-439 <KOR>

A/Cross-references: GB:AE005672; PIDN:AAK75438.1; PID:g14972823; GSPDB:GN00164; TIGR

A/Experimental source: strain TIGR4

C/Genetics:

A/Gene: SPI340

Query Match

Best Local Similarity 24.1%; Pred. No. 0.3; Mismatches 39; Conservative 35; Mismatches 64; Indels 24; Gaps 8;

QY 39 LQKLFARKMKILGTIILGIMFSGVIFLTLKPYRPFIFLSGYPFGSVLPFS 98
 Db 210 LQQIF-QDSTILAKAQIIFNIVILDIYIY---QLSPFIQSCFIKLST-----ITRRNN 260
 QY 99 GAFLIAYKRRKTEETILISIMNLSALRAIAGIILLTFGR-ILDQNYICGYSHQNS-QC 156
 Db 261 FMFTV-----SKWNLAYCKSVIKSISA-AITGAILIISFQMSQNILSQFOQDDSDLEL 313
 QY 157 KATVIVFLGILITMTFSIIEFLISLPFSILGCHSECCDCEQ 198
 Db 314 KVAIVVAGAILIVIANIISI-----APLSSHQRIETIQ 349

Search completed: February 24, 2003, 13:05:09
 Job time : 18.5833 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2003, 12:59:26 ; Search time 18.8707 Seconds

(without alignments)
437.387 Million cell updates/sec

Title: US-09-735-712-2

Perfect score: 1015

Sequence: 1 MDSSTAHSVFLVFPPEIRA.....ISLPFSILGSHMDQCDEQC 199

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt-40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	172	16.9	291	1	CD20_MOUSE
2	168	16.6	297	1	CD20_HUMAN
3	159	15.7	243	1	FCEB_RAT
4	158.5	15.6	235	1	FCEB_MOUSE
5	123.5	12.2	244	1	FCEB_HUMAN
6	95	9.4	408	1	Y917_AQUAE
7	92.5	9.1	458	1	TCR_STRAG
8	92	9.1	324	1	NU1M_DROYA
9	91.5	9.0	666	1	NU5M_CHOCR
10	90	8.9	312	1	NU1M_DROSU
11	87	8.6	312	1	NU1M_DROME
12	87	8.6	382	1	RFC_SHIFL
13	86	8.5	880	1	YE21_ARCFU
14	85.5	8.4	458	1	TCR_BACST
15	85.5	8.4	458	1	TCR_STRPN
16	85	8.4	446	1	NU4M_DROME
17	84.5	8.3	333	1	NU2M_APILI
18	84	8.3	198	1	COP_CLOPE
19	84	8.3	590	1	CAN1_YEAST
20	83	8.2	291	1	YC38_PORPU
21	82	8.1	578	1	AC22_STRCO
22	82	8.1	652	1	NU5M_PODAN
23	82	8.1	669	1	NU5M_ARATH
24	82	8.1	704	1	LCRD_YEREN
25	81.5	8.0	173	1	Y366_RICPR
26	81.5	8.0	252	1	MPU1_DROME
27	81.5	8.0	365	1	CYB_ASCSU
28	81	8.0	664	1	NU5C_SYNP2
29	80	7.9	319	1	YMKR_BACSU
30	80	7.9	397	1	YLMR_CABEL
31	79.5	7.8	241	1	YLSN_HUMAN
32	79.5	7.8	346	1	YAMT_BACAD
33	79.5	7.8	447	1	NU4M_ANOGA

34	79.5	7.8	656	1	YAN9_SCHPO	010074 schizosacch
35	79.5	7.8	657	1	NU5M_EMENT	P11628 emericella
36	79.5	7.8	1368	1	Y064_MYCPN	P75613 mycoplasma
37	79	7.8	228	1	YMCC_EMENT	P03883 emericella
38	79	7.8	324	1	NU1M_SOUAC	Q92554 squallus aca
39	79	7.8	532	1	YABM_BACSU	P37555 bacillus su
40	79	7.8	1873	1	CCAS_RABIT	P07293 oryctolagus
41	78.5	7.7	279	1	YS95_CAREL	010969 caenorhabdi
42	78.5	7.7	409	1	NU4M_ASCSU	P24880 ascaris suu
43	78.5	7.7	956	1	MTN2_MOUSE	008746 mus musculu
44	78	7.7	221	1	ATP6_RHISA	Q99821 rhipicephal
45	78	7.7	347	1	NU2M_HUMAN	P03891 homo sapien

ALIGNMENTS

RESULT 1
CD20_MOUSE STANDARD; PRT; 291 AA.
AC P19437;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE B-cell surface protein CD20 homolog (B-cell differentiation antigen
LY-44).
GN MS4A1 OR CD20 OR LY-44 OR MS4A2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID-10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-89067519; PubMed-2461992;
RA Tedder T.F., Klejman G., Distèche C.M., Adler D.A., Schlossman S.F.,
SA Saito H.;
RT "Cloning of a complementary DNA encoding a new mouse B lymphocyte
differentiation antigen, homologous to the human B1 (CD20) antigen,
and localization of the gene to chromosome 19.";
RL J. Immunol. 141:4388-4394(1988).
[2]
SEQUENCE FROM N.A.
RY STRAIN-C57BL/6J; TISSUE-Thymus;
RX MEDLINE-21085660; PubMed-11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Atzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischman W., Gaasterland T., Gissi C., King B., Kochia H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schmitt L.M., Staudli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsh G.,
RA Blake J., Bonfelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima U., Mazzarelli U., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Wetz C., Whitlaker C., Wilmings L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohetsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -1- FUNCTION: THIS PROTEIN MAY BE INVOLVED IN THE REGULATION OF
CC B-CELL ACTIVATION AND PROLIFERATION.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- PTM: PHOSPHORYLATED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE MS4A FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its


```

RESULT 3
FCER1_RAT
ID FCER1_RAT STANDARD: PRT: 243 AA.
AC P13386;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE High affinity immunoglobulin epsilon receptor beta-subunit (FCER1)
GN (Ige Fc receptor, beta-subunit) (Fc epsilon receptor I beta-chain).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
ON NCBI_TaxID=10116;
RX MEDLINE=89359361; PubMed=2527850;
RA Kinet J.-P., Blank U., Ra C., White K., Metzger H., Kochan J.;
RT "Isolation and characterization of cDNAs coding for the beta subunit
RT of the high-affinity receptor for immunoglobulin E.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:6483-6487(1988).
CC -1- FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULIN EPSILON. HIGH
CC AFFINITY RECEPTOR. RESPONSIBLE FOR INITIATING THE ALLERGIC
CC RESPONSE. BINDING OF ALLERGEN TO RECEPTOR-BOUND IGE LEADS TO CELL
CC ACTIVATION AND THE RELEASE OF MEDIATORS (SUCH AS HISTAMINE)
CC RESPONSIBLE FOR THE MANIFESTATIONS OF ALLERGY. THE SAME RECEPTOR
CC ALSO INDUCES THE SECRETION OF IMPORTANT LYMPHOKINES.
CC -1- SUBUNIT: TETRAMER OF AN ALPHA CHAIN, A BETA CHAIN, AND TWO
CC DISULFIDE LINKED GAMMA CHAINS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE MS4A FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M22923; AAA41149.1; -.
DR PIR: A31231; A31231.
KW Ige-binding protein; Receptor; Transmembrane.
FT DOMAIN 1 59 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 60 79 POTENTIAL.
FT DOMAIN 80 97 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 98 117 POTENTIAL.
FT DOMAIN 118 130 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 131 150 POTENTIAL.
FT DOMAIN 151 179 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 180 199 POTENTIAL.
FT DOMAIN 200 243 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 243 AA; 26730 MW; 471DFA59B6868B5D CRC64;

Query Match 15.7%; Score 159; DB 1; Length 243;
Best Local Similarity 28.0%; Pred. No. 8.3e-07;
Matches 45; Conservative 29; Mismatches 71; Indels 16; Gaps 3;

```

```

ID FCER1_MOUSE STANDARD: PRT: 235 AA.
AC P20490;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE High affinity immunoglobulin epsilon receptor beta-subunit (FCER1)
DE (Ige Fc receptor, beta-subunit) (Fc epsilon receptor I beta-chain).
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxID=10090;
RX MEDLINE=89359361; PubMed=2527850;
RA Ra C., Jouvin M.H.E., Kinet J.-P.;
RT "Complete structure of the mouse mast cell receptor for Ige (Fc
RT epsilon RI) and surface expression of chimeric receptors (rat-mouse-
RT human) on transfected cells.";
RL J. Biol. Chem. 264:15323-15327(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129;
RA Hiraoka S., Watanabe M., Takagaki Y., Fujita-Suzuki K., Shinohara N.,
RA Okumura K., Ra C.;
RT "The genomic structure of the allergy associated Fc receptor beta
RT subunit and its high content of SINDS.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULIN EPSILON. HIGH
CC AFFINITY RECEPTOR. RESPONSIBLE FOR INITIATING THE ALLERGIC
CC RESPONSE. BINDING OF ALLERGEN TO RECEPTOR-BOUND IGE LEADS TO CELL
CC ACTIVATION AND THE RELEASE OF MEDIATORS (SUCH AS HISTAMINE)
CC RESPONSIBLE FOR THE MANIFESTATIONS OF ALLERGY. THE SAME RECEPTOR
CC ALSO INDUCES THE SECRETION OF IMPORTANT LYMPHOKINES.
CC -1- SUBUNIT: TETRAMER OF AN ALPHA CHAIN, A BETA CHAIN, AND TWO
CC DISULFIDE LINKED GAMMA CHAINS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE MS4A FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J05019; AAA37601.1; -.
DR PIR: B34342; B34342.
KW MGD; MGI:95495; Ms4a1.
FT DOMAIN 1 51 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 52 71 POTENTIAL.
FT DOMAIN 72 89 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 90 109 POTENTIAL.
FT DOMAIN 110 122 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 123 142 POTENTIAL.
FT DOMAIN 143 171 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 172 191 POTENTIAL.
FT DOMAIN 192 235 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 235 AA; 1C2D6B195738168 CRC64;

Query Match 15.6%; Score 158.5; DB 1; Length 235;
Best Local Similarity 30.5%; Pred. No. 8.9e-07;
Matches 60; Conservative 29; Mismatches 81; Indels 27; Gaps 9;

```

RESULT 4
FCER1_MOUSE

Y 57 FGIMTFSGVIFETLLKPY-----PRPFIFLSGYPFGSVLFGINGAFILAVKRTTE 111

Db	60	VGLLICLCGCTIVCSVL--YSDPEDEVEVLVLTGLPFWGVAFLVLSGLFSLIISSEKNTL	116
Qy	112	TLIIISRLMNLSSLRALAGILITLTFEFLIDQNTICGYSHNSQCAVY-----VLF	164
Db	117	YLVKSGISGANIVSSIAAGTGMILIT-----NLTNFAYMNN-CNVTEDDGCFVASEP	169
Qy	165	GLVTLMTFFSIIELFIS	181
Db	170	TELVIIMLFLITLAFCS	186
RESULT 5			
FCERB_HUMAN	STANDARD;	PRT;	244 AA.
AC	001362;		
DT	01-APR-1993 (Rel. 25, Created)		
DT	01-APR-1993 (Rel. 25, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	High affinity immunoglobulin epsilon receptor beta-subunit (FCERI)		
DE	(IGE Fc receptor, beta-subunit) (Fc epsilon receptor I beta-chain).		
GN	MS4A2 OR FCERB1B.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiinae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=92316966; PubMed=1535625;		
RA	Kuester H., Zhang L., Brihi A.T., Macglashan D.W., Kinet J.-P.;		
RT	"The gene and cDNA for the human high affinity immunoglobulin E		
RL	receptor beta chain and expression of the complete human receptor.";		
RN	J. Biol. Chem. 267:12782-12787(1992).		
RP	[2]		
RX	SEQUENCE FROM N.A.		
RX	MEDLINE=92339505; PubMed=1386024;		
RA	Maekawa K., Imagawa N., Tanaka Y., Harada S.;		
RT	"Determination of the sequence coding for the beta subunit of the		
RL	human high-affinity IGE receptor.";		
RN	FEBS Lett. 302:161-165(1992).		
RP	[3]		
RX	VARIANT GLY-237.		
RX	MEDLINE=96414302; PubMed=8817330;		
RA	Hill M.R., Cookson W.O.;		
RT	"A new variant of the beta subunit of the high-affinity receptor for		
RL	immunoglobulin E (Fc epsilon RI-beta E237G): associations with		
RN	measures of atopy and bronchial hyper-responsiveness.";		
RP	Hum. Mol. Genet. 5:959-962(1996).		
RP	[4]		
RX	VARIANT GLY-237.		
RX	MEDLINE=96440420; PubMed=8842731;		
RA	Shirakawa T., Mao X.-Q., Sasaki S., Enomoto T., Kawai M., Morimoto K.,		
RT	Hopkin J.;		
RL	"Association between atopic asthma and a coding variant of Fc-epsilon-		
RN	RI-beta in a Japanese population.";		
RP	Hum. Mol. Genet. 5:1129-1130(1996).		
RP	[5]		
RP	ERRATUM.		
RX	MEDLINE=97123518; PubMed=8968765;		
RA	Shirakawa T., Mao X.-Q., Sasaki S., Enomoto T., Kawai M., Morimoto K.,		
RT	Hopkin J.;		
RL	Hum. Mol. Genet. 5:2068-2068(1996).		
CC	-1- FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULINS EPSILON. HIGH		
CC	AFFINITY RECEPTOR. RESPONSIBLE FOR INITIATING THE ALLERGIC		
CC	RESPONSE. BINDING OF ALLERGEN TO RECEPTOR-BOUND IGE LEADS TO CELL		
CC	ACTIVATION AND THE RELEASE OF MEDIATORS (SUCH AS HISTAMINE)		
CC	RESPONSIBLE FOR THE MANIFESTATIONS OF ALLERGY. THE SAME RECEPTOR		
CC	ALSO INDUCES THE SECRETION OF IMPORTANT LYMPHOKINES.		
CC	-1- SUBUNIT: TETRAMER OF AN ALPHA CHAIN, A BETA CHAIN, AND TWO		
CC	DISULFIDE LINKED GAMMA CHAINS.		
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.		
CC	-1- TISSUE SPECIFICITY: FOUND ON THE SURFACE OF MAST CELLS AND		
CC	BASOPHILS.		
CC	-1- POLYMORPHISM: VARIANT GLU-237 HAS BEEN FOUND TO BE PRESENT IN ABOUT		

```

CC      5.3% OF A 1004 INDIVIDUALS POPULATION SAMPLE IN AUSTRALIA. IT
CC      SEEMS TO BE A RISK FACTOR FOR ANOPIIC DERMATITIS AND ASTHMA.
CC      -1- SIMILARITY: BELONGS TO THE MS4A FAMILY.
CC      -----
CC      THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; D10583; BAA01440.1; -.
DR      EMBL; M89796; AAA60269.1; -.
DR      PIR; S21154; S21154.
DR      PIR; A42806; A42806.
DR      Genew; HGNC:7316; MS4A2.
DR      MIM; 147138; -.
KW      19F-binding protein; Receptor; Transmembrane; Polymorphism.
FT      DOMAIN 1 59
FT      TRANSMEM 60 79
FT      DOMAIN 80 97
FT      TRANSMEM 98 117
FT      DOMAIN 118 130
FT      TRANSMEM 131 150
FT      DOMAIN 151 180
FT      TRANSMEM 181 200
FT      DOMAIN 201 244
FT      VARIANT 237
FT      /FTID=VAR_003965
SQ      SEQUENCE 244 AA; 26533 MW; CE523102D5F567AF CRC64;
Query Match 12.2%; Score 123.5; DB 1; Length 244;
Best Local Similarity 25.6%; Pred. No. 0.00089;
Matches 50; Conservative 33; Mismatches 93; Indels 19; Gaps 7;
QY      9 PVFLVPEPEITASYSVESSELSATFTFNSQPG---KLFARMKILGTLQILFGIMTSFG 65
Db      21 PAFEVL--EISPOEVSSGRLLKS--ASSPPLHTWLVYVKKQELGYQIILTMATICCFG 76
QY      66 VILFTLLKPPRPPEFT-FLISGYPFGWSVLFINSGAFLLAVKRRKTEYLILSRINML 123
Db      77 TVVCSVLDISHIEDIIFSSFRAGYPFGALFFSISGMSLIISERNATYIVRSLGANT 136
QY      124 SALALAIAGILLTGFILIDONYICGYSHONS-----CKANTVLELILITIMTF-SIIE 177
Db      137 SSINGGIGITTL---IINLKSILATYTHHSQCKFFETKCFMASFSFEIVMMMLFTYIG 192
QY      178 LFLSLPFSILGCHSE 192
Db      193 LGSVSLTICGAGEE 207
RESULT 6
Y917_AQUAE STANDARD; PRT; 408 AA.
ID Y917_AQUAE
AC 067062;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypoetical protein AQ_917.
GN AQ_917.
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificae (class); Aquificales; Aquificaceae;
OC Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=VF5;
RA MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aulay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;

```


[illegible]

```
CC ----- This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; Z47547; CAAB7625.1; .
DR InterPro; IPR003916; NADHud_oxred5.
DR InterPro; IPR001750; Oxidored_g1.
DR InterPro; IPR001516; Oxidored_g1_N.
DR Pfam; PF00361; oxidored_g1; 1.
DR Pfam; PF00662; oxidored_g1_N; 1.
DR PRINTS; PRO1434; NADHDGNASE5.
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
SQ SEQUENCE 666 AA; 75628 MW; 37E86P2C24B9D360 CRC64;

Query Match          9.0%; Score 91.5; DB 1; Length 666;
Best Local Similarity 21.2%; Pred. No. 1.3;
Matches 56; Conservative 32; Mismatches 83; Indels 93; Gaps 9;

QY      16 PEITSEYESEFELSTATTSTOSPLCKLPARKKILGTLQILEGMTSGVFLEFLTKLP 75
       |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB      394 PLRTGFYSKDPLELTQLTISSSNLOMSYISFRACWLGTVSFV-TSYFSRLIYLFLNMT 452
       ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY      76 -----YPRFFETL-----SGYPFGSVLEFIN----- 98
       |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB      453 NLAKSLMLVHSSLMLFEPLIIISGISIPAGYLIRDLFVGSGSDPGWAALFILPKHSTF 512
       ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY      99 -----GAFLAVRKRTTEILLISRIINLLSALAIAAG----- 131
       ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB      513 IEAEELPIVVKWLPEPLISLGLIFFSFVQIFLKTFEFKSNLNLLSFTFLINKKMWDV 572
       ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY      132 -----TILLTFGR-----TIDQNYI-----CGYSHQNS-----OCKAVTVLELG 165
       ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB      573 LYNRLIVLPILNFSGISFKILDKRGSFIELSGPYGFTKRVSWSQLIKLQIGQITHLYLF 632
       ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY      166 ILITLMFPSITEL--FISLPSITL 187
       ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB      633 MIPTCSFSIILVSYINLTFNL 656
       ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

RESULT 10
NUIM_DROSU
ID NUIM_DROSU STANDARD; PRT; 312 AA.
AC P51937;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 1 (EC 1.6.5.3).
GN MT-NDI OR NDI.
OS Drosophila subobscura (Fruit fly).
OC Mitochondrion.
OG Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
RX NCBI_TaxId=7241;
[1]
RP SEQUENCE OF 1-152 FROM N.A.
RX MADLINE=95106285; PubMed=7528808;
RA Barrio E., Latorre A., Moya A.;
RT "Phylogeny of the Drosophila obscure species group deduced from
RT mitochondrial DNA sequences";
RL J. Mol. Evol. 39:478-488(1994).
[2]
RP SEQUENCE OF 50-312 FROM N.A.
RC STRAIN=SSP. TUE 3;
RA Volz-Lingenhl A.;
RL Submitted (JUN-1994) to the EMBL/GenBank/DBJ databases.
CC -1 CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -1 SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 1 FAMILY.
```


CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL: U07288; AAA76626.1; -
 CC EMBL: X65130; CAA46260.1; -
 CC FLYbase; FBgn0012956; Desub.mtc:ND1.
 CC InterPro; IPR001694; Resp.NADH_dhl.
 CC Pfam; PF00146; NADHdh; 1.
 CC PROSITE; PS00667; COMPLEX1_ND1_1; 1.
 CC PROSITE; PS00668; COMPLEX1_ND1_2; 1.
 CC Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Transmembrane.
 CC SEQUENCE 312 AA; 35877 MW; AFD1B3BC2BCDE68 CRC64;

Query Match 8.9%; Score 90; DB 1; Length 312;

Best Local Similarity 25.8%; Pred. No. 0.82; Mismatches 63; Indels 54; Gaps 9;

Matches 49; Conservative 24; Mismatches 63; Indels 54; Gaps 9;

QY 48 KILGTIQI-----LFGIMTFSEGVIFLFTLKPYPFRP-----FIPLSGY- 87
 Db 30 KYLGVIQIRKGNKVGIMGIPQPCDAIKLFTKEQYPILSNVLSYISPIFSLSFLV 89
 QY 88 ----PFWGSVLFIN-SGALLIAVKRRTTETLLIILSRIM-----LSALAIA- 130
 Db 90 WMCMPFVKLYSFNGLGLEFLCCTSLGYTVVAGWSSNSNYALLGLRAVAQTISYEVS 149
 QY 131 -GIILTFEGFIL-DONYICGSHONSQCAKAVLFLGILITLMTSIIELFISL----- 182
 Db 150 LAILILSFILIGSYIMTYFFFYQ-----IYWFLLILFPALVWVTSIAETNRT 200
 QY 183 PPSILGCHSE 192
 Db 201 PFDFAEGESE 210

RESULT 11

NOIM_DROME

ID NUIM_DROME STANDARD; PRT; 312 AA.

AC P18929;

DT 01-NOV-1990 (Rel. 16, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE NADH-ubiquinone oxidoreductase chain 1 (EC 1.6.5.3).

MT:ND1 OR ND1.

OS Drosophila melanogaster (Fruit fly).

OC Mitochondrion.

CC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;

CC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;

CC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

OX NCB1_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Breitagne;

RX MEDLINE=86212147; PubMed=3130291;

RA Garesse R.;

RT "Drosophila melanogaster mitochondrial DNA: gene organization and
 RT evolutionary considerations.";

RL Genetics 118:649-663(1988).

CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.

CC -1- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 1 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL: M37275; AAA69715.1; -
 CC EMBL: U37541; AAC47823.1; -
 CC PIR; S01191; S01191.
 CC FLYbase; FBgn0013679; mt:ND1.
 CC InterPro; IPR001694; Resp.NADH_dhl.
 CC Pfam; PF00146; NADHdh; 1.
 CC PROSITE; PS00667; COMPLEX1_ND1_1; 1.
 CC PROSITE; PS00668; COMPLEX1_ND1_2; 1.
 CC Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Transmembrane.
 CC SEQUENCE 312 AA; 35910 MW; B2B01BE6213CE722 CRC64;

Query Match 8.6%; Score 87; DB 1; Length 312;

Best Local Similarity 25.8%; Pred. No. 1.5; Mismatches 63; Indels 54; Gaps 9;

Matches 49; Conservative 24; Mismatches 63; Indels 54; Gaps 9;

QY 48 KILGTIQI-----LFGIMTFSEGVIFLFTLKPYPFRP-----FIPLSGY- 87
 Db 30 KYLGVIQIRKGNKVGIMGIPQPCDAIKLFTKEQYPILSNVLSYISPIFSLSFLV 89
 QY 88 ----PFWGSVLFIN-SGALLIAVKRRTTETLLIILSRIM-----LSALAIA- 130
 Db 90 WMCMPFVKLYSFNGLGLEFLCCTSLGYTVVAGWSSNSNYALLGLRAVAQTISYEVS 149
 QY 131 -GIILTFEGFIL-DONYICGSHONSQCAKAVLFLGILITLMTSIIELFISL----- 182
 Db 150 LAILILSFILIGSYIMTYFFFYQ-----IYWFLLILFPALVWVTSIAETNRT 200
 QY 183 PPSILGCHSE 192
 Db 201 PFDFAEGESE 210

RESULT 12

RFC_SHIFL

ID RFC_SHIFL STANDARD; PRT; 382 AA.

AC P37784;

DT 01-OCT-1994 (Rel. 30, Created)

DT 01-OCT-1994 (Rel. 30, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE O-antigen polymerase.

MT:RFC.

OS Shigella flexneri.

CC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

CC Shigella.

OX NCB1_TaxID=623;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Serotype 2A;

RX MEDLINE=94131953; PubMed=7507920;

RA Morona R., Mavris M., Pallarino A., Manning P.A.;

RT "Characterization of the rfc region of Shigella flexneri.";

RL J. Bacteriol. 176:733-747(1994).

CC -1- FUNCTION: MAY LINK THE O-ANTIGEN TETRASACCHARIDE UNITS INTO LONG

CC CHAINS, GIVING RISE TO TYPICAL SMOOTH IPS.

CC -1- PATHWAY: Lipopolysaccharide biosynthesis.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane

CC (potential).

CC -1- SIMILARITY: TO OTHER O-ANTIGEN POLYMERASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: X71970; CAA50774.1; -

CC Lipopolysaccharide biosynthesis; Transmembrane; Inner membrane.

CC TRANSMEM 11 31 POTENTIAL.

CC TRANSMEM 36 56 POTENTIAL.

CC TRANSMEM 61 81 POTENTIAL.


```

FT TRANSMEM 12 33 POTENTIAL.
FT TRANSMEM 81 100 POTENTIAL.
FT TRANSMEM 111 129 POTENTIAL.
FT TRANSMEM 140 162 POTENTIAL.
FT TRANSMEM 165 185 POTENTIAL.
FT TRANSMEM 201 221 POTENTIAL.
FT TRANSMEM 223 240 POTENTIAL.
FT TRANSMEM 256 276 POTENTIAL.
FT TRANSMEM 297 317 POTENTIAL.
FT TRANSMEM 324 344 POTENTIAL.
FT TRANSMEM 346 365 POTENTIAL.
FT TRANSMEM 432 451 POTENTIAL.
SQ SEQUENCE 458 AA; 50119 MW; 116355-1635(1990).

Query Match
Best Local Similarity 24.7%; Score 85.5; DB 1; Length 458;
Matches 43; Conservative 27; Mismatches 69; Indels 35; Gaps 7;

OY 32 TFSSTQSPLOKLFARKMKILGTIOILFGIMTFSGVIFLFLKRPFRPFILSGYPFWG 91
Db 178 TITVPELMKLRKEVRKIKGHFDIK-GILMSYGVF-----MLFTTSYSI-- 223
OY 92 SVLFINGAGLLAVK--RKTEFLIILSRIMNLASLRATAGIIL--TFGFIIDQNYIC 147
Db 224 SFLIVSVLFLFKHRIKRYTDFPDGLGKNIPEFMIGVLCGGIFCTVAGFVSMYPYM 283
OY 148 GYSHONSCKAVTVLFL-----GIIIT---TLMTFSIIEFLSLPF 184
Db 284 KDVHQLSTALIGSVIIFPGTMSVIRYIGIGILVDRGRPLXVNLIGVTLVSSE 337

RESULT 15
TCR_STRPN STANDARD; PRT; 458 AA.
AC P11063: P72219.
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE tetracycline resistance protein.
OS Streptococcus pneumoniae,
OS Bacillus cereus, and
OS Bacillus subtilis.
OG Plasmid pLS1, Plasmid pBC16, Plasmid pHY163PUL, Plasmid pTB19, and
OG Plasmid pNS1981.
OC Bacteria: Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus;
OC NCBI_Taxid=1313, 1396, 1423;
RN 11]
RP SEQUENCE FROM N.A.
RC SPECIES=S.pneumoniae; PLASMID=PLS1;
RX MEDLINE=87226167; PubMed=2438417;
RA Lacke S.A., Lopez P., Greenberg B., Espinosa M.;
RT "Identification and analysis of genes for tetracycline resistance and
RT replication functions in the broad-host-range plasmid pLS1.";
RL J. Mol. Biol. 192:753-765(1986).
RN 12]
RP SEQUENCE FROM N.A.
RC SPECIES=B.cereus; PLASMID=pBC16;
RX MEDLINE=90221899; PubMed=2109312;
RA Palva A., Vildgren G., Simonen M., Rintala H., Laamanen P.;
RT "Nucleotide sequence of the tetracycline resistance gene of pBC16
RT from Bacillus cereus.";
RL Nucleic Acids Res. 18:1635-1635(1990).
RN 13]
RP SEQUENCE FROM N.A.
RC PLASMID=pHY163PUL;
RA Ishida H., Shubahara H.;
RT "New shuttle vectors for Escherichia coli and Bacillus subtilis. III.
RT Nucleotide sequence analysis of tetracycline resistance gene of
RT pAMalpha1 and ori-177.";
RL Jpn. J. Genet. 60:485-496(1985).
RN 14]

```

```

RP SEQUENCE FROM N.A.
RC PLASMID=pHY163PUL;
RX MEDLINE=83129391; PubMed=6186390;
RA Seizer G., Som T., Itoh T., Tomizawa J.;
RT "The origin of replication of plasmid p15A and comparative studies on
RT the nucleotide sequences around the origin of related plasmids.";
RL Cell 32:119-129(1983).
RN 15]
RP SEQUENCE FROM N.A.
RC SPECIES=B.subtilis; PLASMID=pTB19;
RX MEDLINE=92052681; PubMed=1946749;
RA Oskam L., Hillenga D.J., Venema G., Bron S.;
RT "The large Bacillus plasmid pTB19 contains two integrated
RT rolling-circle plasmids carrying mobilization functions.";
RL Plasmid 26:30-39(1991).
RN 16]
RP SEQUENCE FROM N.A.
RC SPECIES=B.subtilis; STRAIN=168 / Marburg; PLASMID=pNS1981;
RX MEDLINE=86287699; PubMed=3090576;
RA Sakauchi R., Shishido K., Hoshino T., Furukawa K.;
RT "The nucleotide sequence of the tetracycline resistance gene of
RT plasmid pNS1981 from Bacillus subtilis differs from pTH15 from a
RT thermophilic Bacillus by two base pairs.";
RL Plasmid 16:72-73(1986).
CC -1- FUNCTION: RESISTANCE TO TETRACYCLINE BY AN ACTIVE TETRACYCLINE
CC EFFLUX. THIS IS AN ENERGY-DEPENDENT PROCESS THAT DECREASES THE
CC ACCUMULATION OF THE ANTIBIOTIC IN WHOLE CELLS. THIS PROTEIN
CC FUNCTIONS AS A METAL-TETRACYCLINE/H+ ANTI-PORTER.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN
CC AS THE DRUG RESISTANCE TRANSLUCASE FAMILY).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X51366; CAA35751.1; -
DR EMBL: M29725; AAA98167.1; -
DR EMBL: D13792; -; NOT_ANNOTATED_CDS.
DR EMBL: M63891; AAA98304.1; -
DR EMBL: D00006; BAA00005.1; -
DR PIR: S09234; YTB506.
DR PIR: G25599; YTSOG.
DR PIR: JQ1211; JQ1211.
DR InterPro: IPR001411; TCR_TetB.
DR PRINTS: PR01036; TCRTEB.
KW Antibiotic resistance; Transmembrane; Transport; Symport; Plasmid.
FT TRANSMEM 12 33 POTENTIAL.
FT TRANSMEM 81 100 POTENTIAL.
FT TRANSMEM 111 129 POTENTIAL.
FT TRANSMEM 140 162 POTENTIAL.
FT TRANSMEM 165 185 POTENTIAL.
FT TRANSMEM 201 221 POTENTIAL.
FT TRANSMEM 223 240 POTENTIAL.
FT TRANSMEM 256 276 POTENTIAL.
FT TRANSMEM 297 317 POTENTIAL.
FT TRANSMEM 324 344 POTENTIAL.
FT TRANSMEM 346 365 POTENTIAL.
FT TRANSMEM 432 451 POTENTIAL.
SQ SEQUENCE 458 AA; 50092 MW; 6A708777D44C2074 CRC64;

Query Match
Best Local Similarity 24.7%; Score 85.5; DB 1; Length 458;
Matches 43; Conservative 27; Mismatches 69; Indels 35; Gaps 7;

OY 32 TFSSTQSPLOKLFARKMKILGTIOILFGIMTFSGVIFLFLKRPFRPFILSGYPFWG 91
Db 178 TITVPELMKLRKEVRKIKGHFDIK-GILMSYGVF-----MLFTTSYSI-- 223

```

Mon Feb 24 13:54:02 2003

us-09-735-712-2.rsp

Page 10

Oy	92	SULFNSASFLIAVK--RKTETIIIRMINLLSARLAGITLL--TFGFLIDQNYIC	147
Db	224	SFLIATVSLFIFPKHIRKTVDPVDDGLKNIDPFMLGIGITFECTVAGFVSMPYIM	283
Oy	148	GYSHNSCCRAVATLFL-----GILL-----TLMTSIIELFSLSPF	184
Db	284	KDVHQLSTALAEIGSVIIFPGTMSVLIIFGIGILVDRNGPLVNLGVTYLSVSEF	337

Search completed: February 24, 2003, 13:03:33
Job time : 20.8707 secs

•

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2003, 13:00:36 ; Search time 30.8793 Seconds
(without alignments)
1327.859 Million cell updates/sec

Title: US-09-735-712-2

Perfect score: 1015

Sequence: 1 MDSTASHPVFLVFPPEITRA.....ISLPISILGCHSEDDCDEQC 199

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP_invertebrate:*
6: SP_mammal:*
7: SP_mhc:*
8: SP_organelle:*
9: SP_phage:*
10: SP_plant:*
11: SP_rodent:*
12: SP_virus:*
13: SP Vertebrate:*
14: SP_unclassified:*
15: SP_rvivirus:*
16: SP_bacteriap:*
17: SP_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1004	98.9	200	4	Q9H3V2
2	998	98.3	200	4	Q9B2H1
3	206.5	20.3	205	4	Q9H3V3
4	206.5	20.3	220	4	Q9B1Y8
5	206.5	20.3	239	4	Q96J05
6	204.5	20.1	197	4	Q9P1S3
7	204.5	20.1	197	4	Q9P1S3
8	190.5	18.8	217	11	Q9D7Z9
9	189.5	18.7	217	11	Q9D7Z9
10	188.5	18.6	214	4	Q96J05
11	184.5	18.2	225	4	Q96J05
12	184.5	18.1	248	4	Q9H2M1
13	183.5	18.1	225	4	Q9H2M1
14	183.5	18.1	248	4	Q9H2M1
15	183.5	18.1	248	4	Q9H2M1
16	183.5	18.1	248	4	Q9H2M1

17	181.5	17.9	225	11	Q99N05	Q99N05 mus musculu
18	181	17.8	226	11	Q9ES61	Q9ES61 mus musculu
19	178	17.5	226	11	Q99N10	Q99N10 mus musculu
20	176.5	17.4	267	4	Q9NXX0	Q9NXX0 homo sapien
21	175	17.2	226	11	Q9D3F6	Q9D3F6 mus musculu
22	174.5	17.2	220	11	Q9DBE5	Q9DBE5 mus musculu
23	173	17.0	679	4	Q96J04	Q96J04 homo sapien
24	172	16.9	178	4	Q9H3V1	Q9H3V1 homo sapien
25	169	16.7	268	11	Q9PQZ0	Q9PQZ0 mus musculu
26	168	16.6	135	6	Q95L74	Q95L74 macaca fasc
27	163.5	16.1	213	11	Q920C4	Q920C4 mus musculu
28	162.5	16.0	240	4	Q9GZM8	Q9GZM8 homo sapien
29	162	16.0	249	4	Q9H2N3	Q9H2N3 homo sapien
30	159.5	15.7	247	11	Q99N07	Q99N07 mus musculu
31	158	15.6	244	11	Q99N09	Q99N09 mus musculu
32	156.5	15.4	247	11	Q9D0X1	Q9D0X1 mus musculu
33	156	15.4	244	11	Q9D8W9	Q9D8W9 mus musculu
34	149	14.7	249	4	Q9H2L1	Q9H2L1 homo sapien
35	148.5	14.6	176	11	Q9DC76	Q9DC76 mus musculu
36	148.5	14.6	250	4	Q9H2A5	Q9H2A5 homo sapien
37	147.5	14.5	250	4	Q9B1Y9	Q9B1Y9 homo sapien
38	142.5	14.0	167	4	Q96EG7	Q96EG7 mus musculu
39	140.5	13.8	124	11	Q99MX6	Q99MX6 mus musculu
40	140.5	13.8	190	11	Q9EOY7	Q9EOY7 mus musculu
41	138	13.6	234	11	Q99N04	Q99N04 mus musculu
42	137.5	13.5	230	11	Q9EOY9	Q9EOY9 mus musculu
43	137	13.5	234	11	Q9H3W1	Q9H3W1 mus musculu
44	121	11.9	125	11	Q912Y5	Q912Y5 mus musculu
45	119.5	11.8	116	11	Q9R046	Q9R046 mus musculu

ALIGNMENTS

RESULT 1	PRELIMINARY:	PRT:	200 AA.
Q9H3V2			
ID	Q9H3V2		
AC	Q9H3V2		
DT	01-MAR-2001 (TREMBLrel. 16, Created)		
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)		
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)		
DE	MS4A5 (MS4A5 protein).		
GN	MS4A5.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=TESTIS;		
RA	Ishibashi K., Sasaki S., Marumo F.;		
RT	"Cloning of three CD20 homolog from human, putative calcium channels.";		
RT	Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.		
RL	[2]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=21295030; PubMed=11401424;		
RA	Liang Y., Tedder T.F.;		
RT	"Identification of a CD20-, Fcpsi1onRbeta-, and FTM4-Related Gene Family: Sixteen New MS4A Family Members Expressed in Human and Mouse.";		
RL	Genomics 72:119-127(2001).		
DR	EMBL; AB013103; BAB18739.1; -		
DR	EMBL; AF237907; AAK37416.1; -		
SO	SEQUENCE 200 AA; 22283 MW; F9282E5D15BC514 CRC64;		
Query Match	98.9%;	Score 1004;	DB 4;
Best Local Similarity	99.0%;	Pred. No. 8.4e-86;	Length 200;
Matches 197;	Conservative	0;	Mismatches 2;
			Indels 0;
			Gaps 0;
QY	1 MDSTASHPVFLVFPPEITASEYESTELSATPTSTOSPLQKLPARKKILGTIQLFGIM 60		
DB	1 MDSTASHPVFLVFPPEITASEYESTELSATPTSTOSPLQKLPARKKILGTIQLFGIM 60		

```

QY 61 TFSFGVIFLTLKPYRPPPIFLSGYPWGSVLFINSGLAVLAVKRTETLILSRIM 120
DB 61 TFSFGVIFLTLKPYRPPPIFLSGYPWGSVLFINSGLAVLAVKRTETLILSRIM 120
QY 121 NLSALRAIAGIILTFEGFILDQNYICGYSHONSQCAKAVLFLGILITLMTFSIIELEF 180
DB 121 NLSALRAIAGIILTFEGFILDQNYICGYSHONSQCAKAVLFLGILITLMTFSIIELEF 180
QY 181 SLFPSILGCHSEDDCEQC 199
DB 181 SLFPSILGCHSEDDCEQC 199

RESULT 2
QY 09B2H1 PRELIMINARY; PRT; 200 AA.
AC 09B2H1;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Testis-expressed transmembrane-4 protein.
GN TETM4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21092614; PubMed=11162526;
RA Hulet M.D., Pagler E., Hornby J.R., Hogarth P.M., Eyre H.J.,
RA Baker E., Crawford J., Sutherland G.R., Ohms S.J., Parish C.R.;
RT "Isolation, tissue distribution, and chromosomal localization of a
RT novel testis-specific human four-transmembrane gene related to CD20
RT and fcepsilonRI-beta."
RL Biochem. Biophys. Res. Commun. 280:374-379(2001).
DR EMBL; AF321127; AAK01641.1; -.
SQ SEQUENCE 200 AA; 22249 MW; 7B282E5D15B25BF4 CRC64;

Query Match
Best Local Similarity 98.3%; Score 998; DB 4; Length 200;
Matches 196; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MDSSTASHPRFLVPPPIITSESESTELSTSTQSPLOKLPARKKILIGTIOILFGIM 60
DB 1 MDSSTASHPRFLVPPPIITSESESTELSTSTQSPLOKLPARKKILIGTIOILFGIM 60
QY 61 TFSFGVIFLTLKPYRPPPIFLSGYPWGSVLFINSGLAVLAVKRTETLILSRIM 120
DB 61 TFSFGVIFLTLKPYRPPPIFLSGYPWGSVLFINSGLAVLAVKRTETLILSRIM 120
QY 121 NLSALRAIAGIILTFEGFILDQNYICGYSHONSQCAKAVLFLGILITLMTFSIIELEF 180
DB 121 NLSALRAIAGIILTFEGFILDQNYICGYSHONSQCAKAVLFLGILITLMTFSIIELEF 180
QY 181 SLFPSILGCHSEDDCEQC 199
DB 181 SLFPSILGCHSEDDCEQC 199

RESULT 3
QY 09H3V3 PRELIMINARY; PRT; 205 AA.
AC 09H3V3;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE MS4A4.
GN MS4A4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21142397; PubMed=11245982;
RA Ishihashi K., Suzuki M., Sasaki S., Imai M.;
RT "Identification of a new multigene four-transmembrane family (MS4A)
RT related to CD20, Htm4 and beta subunit of the high-affinity Ige
RT receptor."
RL Gene 264:87-93(2001).
DR EMBL; AB013102; BAB18738.1; -.
SQ SEQUENCE 205 AA; 21790 MW; FDB1899FCC51B3 CRC64;

Query Match
Best Local Similarity 20.3%; Score 206.5; DB 4; Length 205;
Matches 49; Conservative 32; Mismatches 70; Indels 3; Gaps 2;

QY 39 LQKLFAR-KMKILGTIOILFEGIMTFSFGVIFLTLKPYRPPPIFLSGYPWGSVLEFIN 97
DB 20 LQKLFAR-KMKILGTIOILFEGIMTFSFGVIFLTLKPYRPPPIFLSGYPWGSVLEFIN 97
QY 98 SGAFILAVKRTETLILSRIMNLSALRAIAGIILTF--GFILDQNYICGYSHONSQ 155
DB 80 SGSLSTIAGIRTKGLVRSGLGMINITSSVLAASGLINFTSLAFYFHHPRYCNYSNN 139
QY 156 CKAIVLFLGILITLMTFSIIELEFSLPSIIGC 189
DB 140 CHGTMSILMGDGVLLSVLEFCIAVSLSARFC 173

RESULT 4
QY 09B18 PRELIMINARY; PRT; 220 AA.
AC 09B18;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE MS4A4A protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=PLACENTA;
RX MEDLINE=21295030; PubMed=11401424;
RA Liang Y., Tedder T.F.;
RT "Identification of a cd20-, fcepsilonRI-beta-, and htm4-related gene
RT family: sixteen new ms4a family members expressed in human and
RT mouse."
RL Genomics 72:119-127(2001).
DR EMBL; AF237912; AAK37594.1; -.
SQ SEQUENCE 220 AA; 23220 MW; 7648C762197C1179 CRC64;

Query Match
Best Local Similarity 20.3%; Score 206.5; DB 4; Length 220;
Matches 49; Conservative 32; Mismatches 70; Indels 3; Gaps 2;

QY 39 LQKLFAR-KMKILGTIOILFEGIMTFSFGVIFLTLKPYRPPPIFLSGYPWGSVLEFIN 97
DB 35 LQKLFAR-KMKILGTIOILFEGIMTFSFGVIFLTLKPYRPPPIFLSGYPWGSVLEFIN 94
QY 98 SGAFILAVKRTETLILSRIMNLSALRAIAGIILTF--GFILDQNYICGYSHONSQ 155
DB 95 SGSLSTIAGIRTKGLVRSGLGMINITSSVLAASGLINFTSLAFYFHHPRYCNYSNN 154
QY 156 CKAIVLFLGILITLMTFSIIELEFSLPSIIGC 189
DB 155 CHGTMSILMGDGVLLSVLEFCIAVSLSARFC 188

RESULT 5
QY 096J05 PRELIMINARY; PRT; 239 AA.
ID 096J05

```

AC 096305;
 DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
 DE mRNA similar to fc epsilon receptor beta subunit, complete cds
 DE (Membrane-spanning 4-domains, subfamily A, member 4A).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN NCB1_TaxID=9606;
 RP SEQUENCE FROM N.A.
 RC TISSUE=BLOOD;
 RA Fujii Y., Takayama K.;
 RT "Characterization of a fc epsilon receptor beta subunit homolog
 RT lacking ITAM motif";
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 RN [1]
 RC SEQUENCE FROM N.A.
 RP TISSUE=PLACENTA;
 RA Strausberg R.;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB022821; BAB61018.1; -
 DR EMBL; BC020648; AAH20648.1; -
 KM Receptor.
 SQ SEQUENCE 239 AA; 25441 MW; F0617D94669CB447 CRC64;

Query Match 20.3%; Score 206.5; DB 4; Length 239;
 Best Local Similarity 31.8%; Pred. No. 1.5e-11;
 Matches 49; Conservative 32; Mismatches 70; Indels 3; Gaps 2;

QY 39 LOKLFAR-KMKILGTIOILFGIMTFSGVIFLFTLKPYRPFPIFLSGYFPGSVYFIN 97
 DB 54 LQKFLKGEKVLGVVOILALMSLSMGITMCMASNTYGSNPISYIGITMGSVFII 113
 QY 98 SGAFIAVKKRTETETLIIISRMNLSALRAIAGIILTF--GTILDQNYICGYSHQNSQ 155
 DB 114 SGLSIAAGIRTKGLVRSIGMNTSSVLAAGIILNFTSLAFYFHHYCNVYGSNN 173
 QY 156 KRAVYVFLGILITLMTFSLIEFLISLPSTLGC 189
 DB 174 CHGTMSILMGLDGMVLLSVLEFCIAVLSAFGC 207

RESULT 6
 Q9P1S3 PRELIMINARY; PRT; 197 AA.
 AC Q9P1S3;
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
 DE HDOM31P.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhao Z., Huang X., Li N., Zhu X., Cao X.;
 RT "A novel gene from human dendritic cell";
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF068288; AAF65507.1; -
 SQ SEQUENCE 197 AA; 20909 MW; 11F056CA869616A7 CRC64;

Query Match 20.1%; Score 204.5; DB 4; Length 197;
 Best Local Similarity 31.8%; Pred. No. 2e-11;
 Matches 49; Conservative 32; Mismatches 70; Indels 3; Gaps 2;

QY 39 LOKLFAR-KMKILGTIOILFGIMTFSGVIFLFTLKPYRPFPIFLSGYFPGSVYFIN 97
 DB 12 LQKFLKGEKVLGVVOILALMSLSMGITMCMASNTYGSNPISYIGITMGSVFII 71
 QY 98 SGAFIAVKKRTETETLIIISRMNLSALRAIAGIILTF--GTILDQNYICGYSHQNSQ 155

DB 72 SGLSIAAGIRTKGLVRSIGMNTSSVLAAGIILNFTSLAFYFHHYCNVYGSNN 131
 QY 156 KRAVYVFLGILITLMTFSLIEFLISLPSTLGC 189
 DB 132 CHGTMSILMGLDGMVLLSVLEFCIAVLSAFGC 165

RESULT 7
 Q8TEZ6 PRELIMINARY; PRT; 197 AA.
 AC Q8TEZ6;
 DT 01-JUN-2002 (TReMBLrel. 21, Created)
 DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE Four-span transmembrane protein 1.
 GN ASPAN1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hulbert M.D., Pagler E., Hogarth M., Eyre H., Baker E., Crawford J.,
 RA Sutherland G.R., Parish C.R.;
 RT "Isolation of a family of hematopoietic-expressed four-transmembrane
 RT genes related to CD20 and Fc epsilonRI beta";
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF350500; AAL56220.1; -
 KM Transmembrane.
 SQ SEQUENCE 197 AA; 20939 MW; E1EC2B3BF303BA CRC64;

Query Match 20.1%; Score 204.5; DB 4; Length 197;
 Best Local Similarity 31.8%; Pred. No. 2e-11;
 Matches 49; Conservative 31; Mismatches 71; Indels 3; Gaps 2;

QY 39 LOKLFAR-KMKILGTIOILFGIMTFSGVIFLFTLKPYRPFPIFLSGYFPGSVYFIN 97
 DB 12 LQKFLKGEKVLGVVOILALMSLSMGITMCMASNTYGSNPISYIGITMGSVFII 71
 QY 98 SGAFIAVKKRTETETLIIISRMNLSALRAIAGIILTF--GTILDQNYICGYSHQNSQ 155
 DB 72 SGLSIAAGIRTKGLVRSIGMNTSSVLAAGIILNFTSLAFYFHHYCNVYGSNN 131
 QY 156 KRAVYVFLGILITLMTFSLIEFLISLPSTLGC 189
 DB 132 CHGTMSILMGLDGMVLLSVLEFCIAVLSAFGC 165

RESULT 8
 Q9D7Z9 PRELIMINARY; PRT; 217 AA.
 AC Q9D7Z9;
 DT 01-JUN-2001 (TReMBLrel. 17, Created)
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
 DE 2200009H22Rik protein.
 GN Mus musculus (Mouse).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=STOMACH;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai T., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Glisli C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,

RA Schiml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher M., Fujita M., Gariboldi M.,
 RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamaya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Rilling B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schenbach C., Seta T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.,
 RT "Functional annotation of a full-length mouse cDNA collection."
 RT EMBL: AK006552; BAB25808.1; -
 DR MGD: MG1:1920906; 2200009H22R1K.
 SQ SEQUENCE 217 AA; 23643 MW; 2B6C4FF0287B543F CRC64;

Query Match 18.8%; Score 190.5; DB 11; Length 217;
 Best Local Similarity 26.4%; Pred. No. 4.4e-10;
 Matches 55; Conservative 41; Mismatches 79; Indels 33; Gaps 5;

OY 2 DSSTAHPVFLVPEPTTASEYESTELSATTFSTQSPLOKLFARKKILGTIQLFGIMT 61
 DB ERTTRKSPGNGINPDK-----DESGPTQQRQDCLKRLKAEIKVAIOMCAVY 59
 OY 62 FSEGVIFLTLKPYRPFPIF---LSGYPMGVSFLFNSGAFILAVKRTETLII 117
 DB LALGIIILASG--PPVPYFNSVFSVLKSGPFGALFGLFASGLISITTEKSTKLVDAS 117
 OY 118 RIMNLSALRAIAGIILTFGGF-----LDONTICG-YSHONSCKAV 159
 DB 118 LFTNLSVSPAFVGIITISVSLGLHPASEQCKSKELSIEDHYQPPYNSDRSCAVT 177
 OY 160 TVFLGLITLMTFSIIEFLSPFSIL 187
 DB 178 KSLTGLALSVMILISVLEGLALSLAML 205

RESULT 9
 O99N08 PRELIMINARY; PRT; 217 AA.
 AC O99N08:
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE MS4A6C protein.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 NC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=TOTAL FETUS;
 RX MEDLINE=21295030; PubMed=11401424;
 RA Liang Y., Tedder T.F.;
 RT "Identification of a cd20-, fcepsilonibeta-, and htm4-related gene
 RT family: sixteen new ms4a family members expressed in human and
 RT mouse";
 RL Genomics 72:119-127(2001).
 DR EMBL: AF237910; AAK37419.1; -
 SQ SEQUENCE 217 AA; 23622 MW; 8BF521AF22DBB7BD CRC64;

Query Match 18.7%; Score 189.5; DB 11; Length 217;
 Best Local Similarity 26.0%; Pred. No. 5.4e-10;
 Matches 54; Conservative 42; Mismatches 79; Indels 33; Gaps 5;

OY 2 DSSTAHPVFLVPEPTTASEYESTELSATTFSTQSPLOKLFARKKILGTIQLFGIMT 61
 DB ERTTRKSPGNGINPDK-----DESGPTQQRQDCLKRLKAEIKVAIOMCAVY 59
 OY 62 FSEGVIFLTLKPYRPFPIF---LSGYPMGVSFLFNSGAFILAVKRTETLII 117
 DB LALGIIILASG--PPVPYFNSVFSVLKSGPFGALFGLFASGLISITTEKSTKLVDAS 117

OY 118 RIMNLSALRAIAGIILTFGGF-----LDONTICG-YSHONSCKAV 159
 DB 118 LFTNLSVSPAFVGIITISVSLGLHPASEQCKSKELSIEDHYQPPYNSDRSCAVT 177
 OY 160 TVFLGLITLMTFSIIEFLSPFSIL 187
 DB 178 KSLTGLALSVMILISVLEGLALSLAML 205

RESULT 10
 O96HJ5 PRELIMINARY; PRT; 214 AA.
 AC O96HJ5:
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Unknown (protein for MGC:14809).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 NC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LUNG;
 RA Strausberg R.;
 RL Submitted (May-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: BC008487; AA08487.1; -
 SQ SEQUENCE 214 AA; 22933 MW; D30C276213DBBF3D CRC64;

Query Match 18.6%; Score 188.5; DB 4; Length 214;
 Best Local Similarity 28.7%; Pred. No. 6.6e-10;
 Matches 60; Conservative 40; Mismatches 84; Indels 25; Gaps 8;

OY 1 MDSSTAH-SPVFLVPEPTTASEYESTELSATTFSTQSPLOKLFARKKILGTIQLFGI 59
 DB 11 LGSASAHGTPGSEAGEELNTSVQPIDG-----PDYGR--AKLQVLAIOILNA 60
 OY 60 MTFSEGVIFLTLKPYRPFPIF---LSGYPMGVSFLFNSGAFILAVKRTETLII 115
 DB 61 MILALGV-FLGSLQYRHHQKHFPEFTYTGPIGCAVFFCCSGLSVAGIKPRTWIQ 119
 OY 116 LSRIMNLSALRAIAGIILTFGGFIDONTY--CGYSHONSQ---CKAVVFLGILIT 169
 DB 120 NSFGMNIAATAIATVGTAFSLINAVNIQSLKSC--HSSSPDLCKNMGISNGWSL 176
 OY 170 LMTFSIIEFLSPFSILGCHSEDCCEQ 198
 DB 177 LTLTLELCVITISTIAMCNANCNSRE 205

RESULT 11
 O96PG6 PRELIMINARY; PRT; 225 AA.
 AC O96PG6:
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE MS4A6A.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 NC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RX MEDLINE=21382044; PubMed=11486273;
 RA Liang Y., Buckley T.R., Tu L., Langdon S.D., Tedder T.F.;
 RT "Structural organization of the human MS4A gene cluster on Chromosome
 RT 14q12";
 RL Immunogenetics 53:357-368(2001).
 DR EMBL: AF354930; AAL07357.1; -
 SQ SEQUENCE 225 AA; 24331 MW; 12BF5DCDC71B56A CRC64;

Db 1 MTSQVPNETIIVLPSNVI--NFSQAEKPEPTNGODSLKKHLHAEIKVIGTIQILCGMM 58
QY 61 TFSQGVIFLFLKPK-YRFPPIFL-SCYPWGSVLFINSQAFLLIYVKKRTETLILSR 118
Db 59 VLSLGIILIASASFSPNFQVSTLNSAVPFIQPFIISSLSIATERRLSKLIVHSSL 118
QY 119 IMNLISALRAIAGIILL-----TFGFIIDON-----YICGYSHON--SQCKAV 159
Db 119 VGSIIISALSALVGFIIISVKQATLNPAISOCELDKNNIPTRSYVSFYHDSLYTTDCYTA 178
QY 160 TVLFGIILITMTFSIIIEFLISLPSIL 187
Db 179 KASLAGTSLMLICTLLEFCIAVLTAVL 206

RESULT 15

Q8TEZ5 PRELIMINARY; PRT; 248 AA.
ID Q8TEZ5
C 01-JUN-2002 (TREMblrel. 21, Created)
DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Four-span transmembrane protein 3.1.
GN 4SPAN3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Hullett M.D., Pagler E., Hogarth M., Eyre H., Baker E., Crawford J.,
RA Sutherland G.R., Parish C.R.;
RT "Isolation of a family of hematopoietic-expressed four-transmembrane
RT genes related to CD20 and FcepsilonRI-beta."
RL Submitted (FEb-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF350502; AAL56222.1; .
KW Transmembrane.
SQ SEQUENCE 248 AA; 27026 MW; CB68780BBC0480CD CRC64;

Query Match

18.1%; Score 183.5; DB 4; Length 248;

Best Local Similarity 26.4%; Pred. No. 2.2e-09;

Matches 53; Conservative 40; Mismatches 90; Indels 23; Gaps 6;

QY 1 MDSSTAHSVPVLPPEPTASEYESTELSATTFSTOSPLQKLFARKMKILGTIQLFGIM 60
Db 1 MTSQVPNETIIVLPSNVI--NFSQAEKPEPTNGODSLKKHLHAEIKVIGTIQILCGMM 58
QY 61 TFSQGVIFLFLKPK-YRFPPIFL-SCYPWGSVLFINSQAFLLIYVKKRTETLILSR 118
Db 59 VLSLGIILIASASFSPNFQVSTLNSAVPFIQPFIISSLSIATERRLSKLIVHSSL 118
QY 119 IMNLISALRAIAGIILL-----TFGFIIDON-----YICGYSHON--SQCKAV 159
Db 119 VGSIIISALSALVGFIIISVKQATLNPAISOCELDKNNIPTRSYVSFYHDSLYTTDCYTA 178
QY 160 TVLFGIILITMTFSIIIEFLISLPSIL 187
Db 179 KASLAGTSLMLICTLLEFCIAVLTAVL 206

Search completed: February 24, 2003, 13:04:33
Job time : 32.8793 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2003, 12:59:26 ; Search time 28.6868 Seconds
(without alignments)
692.108 Million cell updates/sec

Title: US-09-735-712-8

Perfect score: 771

Sequence: 1 MDSTAHSPVFLVFPEPTEA.....SLPESILGCHSEDCDEQCQC 149

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A.GeneSeq-101002.*
1: /SID52/gcgdata/geneSeq/geneSeqp-emb1/AA1980.DAT:*
2: /SID52/gcgdata/geneSeq/geneSeqp-emb1/AA1981.DAT:*
3: /SID52/gcgdata/geneSeq/geneSeqp-emb1/AA1982.DAT:*
4: /SID52/gcgdata/geneSeq/geneSeqp-emb1/AA1983.DAT:*
5: /SID52/gcgdata/geneSeq/geneSeqp-emb1/AA1984.DAT:*
6: /SID52/gcgdata/geneSeq/geneSeqp-emb1/AA1985.DAT:*
7: /SID52/gcgdata/geneSeq/geneSeqp-emb1/AA1986.DAT:*
8: /SID52/gcgdata/geneSeq/geneSeqp-emb1/AA1987.DAT:*
9: /SID52/gcgdata/geneSeq/geneSeqp-emb1/AA1988.DAT:*
10: /SID52/gcgdata/geneSeq/geneSeqp-emb1/AA1989.DAT:*
11: /SID52/gcgdata/geneSeq/geneSeqp-emb1/AA1990.DAT:*
12: /SID52/gcgdata/geneSeq/geneSeqp-emb1/AA1991.DAT:*
13: /SID52/gcgdata/geneSeq/geneSeqp-emb1/AA1992.DAT:*
14: /SID52/gcgdata/geneSeq/geneSeqp-emb1/AA1993.DAT:*
15: /SID52/gcgdata/geneSeq/geneSeqp-emb1/AA1994.DAT:*
16: /SID52/gcgdata/geneSeq/geneSeqp-emb1/AA1995.DAT:*
17: /SID52/gcgdata/geneSeq/geneSeqp-emb1/AA1996.DAT:*
18: /SID52/gcgdata/geneSeq/geneSeqp-emb1/AA1997.DAT:*
19: /SID52/gcgdata/geneSeq/geneSeqp-emb1/AA1998.DAT:*
20: /SID52/gcgdata/geneSeq/geneSeqp-emb1/AA1999.DAT:*
21: /SID52/gcgdata/geneSeq/geneSeqp-emb1/AA2000.DAT:*
22: /SID52/gcgdata/geneSeq/geneSeqp-emb1/AA2001.DAT:*
23: /SID52/gcgdata/geneSeq/geneSeqp-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	771	100.0	149	22	AA001213	Novel human membra
2	735.5	95.4	200	22	AAE13062	Human CD20/IGF-rec
3	735.5	95.4	200	22	AA689138	Human secreted pro
4	726.5	94.2	199	22	AA001210	Novel human membra
5	645.5	83.7	201	22	AA10917	Human gene 9 encod
6	645.5	83.7	212	22	AB95818	Human testicular a
7	645.5	83.7	212	22	AA95114	Human reproductive
8	588	76.3	158	22	AB12234	Human secreted pro
9	576	74.7	138	22	AA001212	Novel human membra
10	294	38.1	77	21	AA03074	Human secreted pro

11	263.5	34.2	67	22	AA689142	Human secreted pro
12	221	28.7	81	22	AA001211	Novel human membra
13	161.5	20.9	167	21	AA53632	A bone marrow secr
14	150.5	19.5	214	17	AAW06503	Htm4 protein. Hom
15	150.5	19.5	214	19	AAW41056	Dendritic cell. Hom
16	141.5	18.4	178	22	AAE12073	Human inflammatory
17	139	18.0	225	21	AA94449	Human cell surface
18	139	18.0	225	22	AA655272	Human secreted pro
19	139	18.0	242	21	AA73495	Human secreted pro
20	139	18.0	248	20	AA15225	Human receptor pro
21	139	18.0	248	21	AA91531	Human secreted pro
22	139	18.0	248	22	AAE12072	Dendritic cell (DC
23	139	18.0	248	22	AA670489	Human hHAIRBs-iso
24	139	18.0	248	23	AA690341	Human polypeptide
25	139	18.0	250	20	AA48505	Human breast tumor
26	139	18.0	257	21	AA58419	Lung cancer associ
27	139	18.0	273	21	AA91680	Human secreted pro
28	139	18.0	273	22	AB11989	Human secreted pro
29	139	18.0	273	22	AA625809	Human protein sequ
30	139	18.0	273	22	AA675619	Human colon cancer
31	137.5	17.8	226	22	AA698706	Chandra, a helper
32	136	17.6	248	20	AA136046	Extended human sec
33	134	17.4	220	23	AB690317	Human polypeptide
34	134	17.4	239	20	AA696745	High affinity immu
35	134	17.4	239	21	AA50174	Human high affinity
36	134	17.4	245	21	AA94973	Human secreted pro
37	133.5	17.3	227	22	AB619236	Novel human diagno
38	132	17.1	220	22	AA693512	Human polypeptide
39	131.5	17.1	239	22	AA674584	Human HAIRBs iso
40	131.5	17.1	249	22	AA674583	Human HAIRBs iso
41	129	16.7	133	22	AB12201	Human IGERB homolo
42	128.5	16.7	244	22	AA638825	Immunoglobulin hlg
43	128	16.6	267	20	AA15224	Human receptor pro
44	128	16.6	273	22	AB617004	Novel human diagno
45	128	16.6	299	21	AA91352	Human secreted pro

ALIGNMENTS

RESULT 1
AA001213 standard; Protein: 149 AA.
XX
AC AA001213:
XX
DT 26-SEP-2001 (first entry)
XX
DE Novel human membrane protein #4.
XX
KW Human: membrane protein; membrane receptor; IGE receptor; CD20;
KW physiological disorder.
XX
OS Homo sapiens.
XX
PN WO200146417-A2.
XX
PD 28-JUN-2001.
XX
PE 12-DEC-2000; 2000MO-US33742.
XX
PR 22-DEC-1999; 99US-0171567.
XX
XX (LEXI-) LEXICON GENETICS INC.
XX
XX Walke DW, Turner CA;
XX
XX WPI, 2001-408646/43.
XX
XX N-PSDB; AAS04282.
XX
XX Polynucleotide encoding novel human membrane protein, useful for
XX identifying agonist, antagonist or modifiers or for producing
XX PT antibodies useful in therapeutic, diagnostic and pharmacogenomic

PT applications -
XX
XX Claim 4; Page 31; 32pp; English.
XX
CC The present sequence represents novel human membrane protein #4.
CC Human membrane protein #4 is 1 of 4 human membrane proteins
CC (AAU0210-AAU0213) given in the present invention. These membrane
CC proteins share structural similarity with membrane receptors such as
CC the Ige receptor and mammalian CD20. The novel human membrane proteins
CC are useful for identifying agonists, antagonists and modulators of the
CC membrane proteins, and for producing antibodies specific to the
CC membrane proteins. The membrane proteins can be used for diagnosis,
CC drug screening, pharmacogenomic applications, clinical trial monitoring
CC and the treatment of physiological disorders and diseases. The
CC polynucleotides encoding the membrane proteins can be used to generate
CC PCR primers or probes to identify mutations associated with a particular
CC disease.
XX
SQ Sequence 149 AA;
Query Match 100.0%; Score 771; DB 22; Length 149;
Best Local Similarity 100.0%; Pred. No. 8.9e-86;
Matches 149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDSSTASHPVFLVFPPEITASEYESTELSATTTSTOSPLOKLPARKKIIIGTIOILFGIM 60
DB 1 MDSSTASHPVFLVFPPEITASEYESTELSATTTSTOSPLOKLPARKKIIIGTIOILFGIM 60
QY 61 TFSFGVIFLFTLLKPYRPFIFLSGYPFMGSVLFINSAGFLAVARKKTTETIGITITLM 120
DB 61 TFSFGVIFLFTLLKPYRPFIFLSGYPFMGSVLFINSAGFLAVARKKTTETIGITITLM 120
QY 121 TFSITIEFLISLPSIILGCHSEDCDCEQCC 149
DB 121 TFSITIEFLISLPSIILGCHSEDCDCEQCC 149
RESULT 2
AAEL3062
ID AAEL3062 standard; Protein; 200 AA.
XX
AC AAEL3062;
XX
DT 28-JAN-2002 (first entry)
DE Human CD20/Ige-receptor like protein, asp-96614-al.
XX
XX Human; CD20/Ige-receptor like protein; immunoglobulin E; asp-96614-al;
XX asp-69406-al; cancer; abnormal cell proliferation; autoimmune disease;
XX ovarian cancer; brain cancer; arteriosclerosis; vascular restenosis;
XX rheumatoid arthritis; multiple sclerosis; allergy; dermatitis; asthma;
XX reproductive disease; diabetes; transplant rejection; endometriosis;
XX infertility; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200174903-A2.
XX
PD 11-OCT-2001.
XX
PF 29-MAR-2001; 2001WO-US10048.
XX
PR 30-MAR-2000; 2000US-193728P.
XX
PR 27-NOV-2000; 2000US-0723258.
XX
PA (AMGE-) AMGEN INC.
XX
PI Welcher AA, Calzone FJ;
XX
XX WPI; 2001-662968/76.
XX
XX N-PSDB; AAD21441.
XX
PT Novel CD20/Ige-receptor like polypeptides and polynucleotides,

PT antagonists and antibodies of the polypeptide useful for treating
PT ameliorating or preventing diseases associated with the polypeptide
PT e.g. cancer, asthma -
XX
XX
PS Claim 13; Fig 1; 145pp; English.
XX
CC The invention relates to human CD20/immunoglobulin E (Ige)-receptor
CC like polypeptides designated as asp-96614-al and asp-69406-al and
CC nucleic acid molecules encoding such polypeptides. Polypeptides of
CC the invention are useful for treating, preventing or ameliorating
CC a disease, condition, or disorder which includes cancer such as
CC brain cancer, ovarian cancer; abnormal cell proliferation such as
CC arteriosclerosis, vascular restenosis; pathology from allergens
CC such as allergies, asthma, dermatitis; dysfunction of immune system
CC such as rheumatoid arthritis, autoimmune disease, multiple sclerosis,
CC diabetes, transplant rejection and reproductive diseases such as
CC infertility, preterm labour and delivery, endometriosis etc. They
CC are also useful for identifying antagonists and as immunogens, for
CC raising antibodies which may also be used to prevent, treat or
CC diagnose a number of diseases and disorders. Polynucleotides of the
CC invention are used to map the location of CD20/Ige-receptor like
CC gene and related genes on chromosomes and as hybridisation probes.
CC They are also useful in gene therapy. The present sequence is
CC human CD20/Ige-receptor like protein, asp-96614-al.
XX
SQ Sequence 200 AA;
Query Match 95.4%; Score 735.5; DB 22; Length 200;
Best Local Similarity 74.5%; Pred. No. 2.7e-81;
Matches 149; Conservative 0; Mismatches 0; Indels 51; Gaps 1;
QY 1 MDSSTASHPVFLVFPPEITASEYESTELSATTTSTOSPLOKLPARKKIIIGTIOILFGIM 60
DB 1 MDSSTASHPVFLVFPPEITASEYESTELSATTTSTOSPLOKLPARKKIIIGTIOILFGIM 60
QY 61 TFSFGVIFLFTLLKPYRPFIFLSGYPFMGSVLFINSAGFLAVARKKTTETIGITITLM 112
DB 61 TFSFGVIFLFTLLKPYRPFIFLSGYPFMGSVLFINSAGFLAVARKKTTETIGITITLM 120
QY 113 -----LGITITLMFSITIEFLI 129
DB 121 NFLSALGATAGIILNFGFIIDQNYICGYSHQNSQCKAVTVLFLGITITLMFSITIEFLI 180
QY 130 SLPFSIILGCHSEDCDCEQCC 149
DB 181 SLPFSIILGCHSEDCDCEQCC 200
RESULT 3
AAG89138
ID AAG89138 standard; Protein; 200 AA.
XX
AC AAG89138;
XX
DT 11-SEP-2001 (first entry)
DE Human secreted protein, SEQ ID NO: 258.
XX
XX Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;
XX GENSET.
XX
OS Homo sapiens.
XX
PN WO200142451-A2.
XX
PD 14-JUN-2001.
XX
PF 07-DEC-2000; 2000WO-IB01938.
XX
XX 08-DEC-1999; 99US-0169629.
XX
XX 06-MAR-2000; 2000US-0187470.
XX
PA (GENSET) GENSET.

XX Dumas Milne Edwards J, Bougueleret L, Jobert S;
PI WPI: 2001-367870/38.
DR N-PSDB; AAH64741.
XX
PT Full length GENSET human nucleic acids encoding potentially secreted
PT proteins; useful in gene therapy and vaccination against a variety of
PT diseases, and for diagnosis of those diseases -
XX
PS Claim 21, Page 802-803; 921pp; English.
XX
CC The invention relates to full length GENSET human nucleic acids encoding
CC potentially secreted proteins. The nucleic acids and the polypeptides of
CC they encode may be used in the prevention, treatment and diagnosis of
CC diseases associated with inappropriate GENSET gene expression. For
CC example, they be used to treat disorders associated with decreased
CC GENSET gene expression by rectifying mutations or deletions in a
CC patient's genome that affect the activity of GENSET or by supplementing
CC the patient's own production of GENSET polypeptides. Conversely,
CC antisense nucleic acid molecules may be administered to down regulate
CC GENSET expression by binding with the cells' own genes and preventing
CC their expression. The sense and antisense nucleic acids may also be
CC used as DNA probes in diagnostic assays to detect and quantitate the
CC presence of similar nucleic acid sequences in samples, and hence to
CC determine which patients may be in need of restorative therapy.
CC The GENSET polypeptides may be used as antigens in the production of
CC antibodies and in assays to identify modulators (agonists and
CC antagonists) of GENSET polypeptide expression and activity. The
CC present sequence is a GENSET polypeptide of the invention.
XX
SQ Sequence 200 AA;
Query Match 95.4%; Score 735.5; DB 22; Length 200;
Best Local Similarity 74.3%; Pred. No. 2,7e-81;
Matches 149; Conservative 0; Mismatches 0; Indels 51; Gaps 1;
QY 1 MDSSTASPVFLVPPPTITASEVSESTELASTTSTOSPLOKLPARKKIIIGTIOILGIM 60
Db 1 MDSSTASPVFLVPPPTITASEVSESTELASTTSTOSPLOKLPARKKIIIGTIOILGIM 60
QY 61 TFSFGVIFLFTLKPYRPFPIFLSGYFPGSVLFINSGAFLIAVKRKTET----- 112
Db 61 TFSFGVIFLFTLKPYRPFPIFLSGYFPGSVLFINSGAFLIAVKRKTETLILSRIM 120
QY 113 -----LGIITITMTSIIIEFT 129
Db 121 NLSALRAIAGIILLTFEGTILDONYIGYSHONSOKRAVTVLFGIITITMTSIIIEFT 180
QY 130 SLPPSIIIGCHSEDDCCGCC 149
Db 181 SLPPSIIIGCHSEDDCCGCC 200
RESULT 4
AAU01210
ID AAU01210 standard; protein; 199 AA.
XX AAU01210;
AC
XX 26-SEP-2001 (first entry)
DT
XX
XX Novel human membrane protein #1.
DE
XX
XX Human; membrane protein; membrane receptor; Ige receptor; CD20;
KW physiological disorder.
XX
XX Homo sapiens.
OS
XX
XX
FH Key Location/Qualifiers
FT Misc-difference 199 /note="Encoded by TGTGTG"
XX

PN WO200146417-A2.
XX
XX 28-JUN-2001.
PD
XX
XX 12-DEC-2000; 2000WO-US33742.
PF
XX
XX 22-DEC-1999; 99US-0171567.
PR
XX
XX (LEXI-) LEXICON GENETICS INC.
PA
XX
XX Walke DW, Turner CA;
PI
XX
XX WPI: 2001-408646/43.
DR N-PSDB; AAS04279.
XX
XX Polynucleotide encoding novel human membrane protein, useful for
PT identifying agonist, antagonist or modifiers or for producing
PT antibodies useful in therapeutic, diagnostic and pharmacogenomic
PT applications -
XX
PS Claim 2; Page 29-30; 32pp; English.
XX
XX The present sequence represents novel human membrane protein #1.
CC Human membrane protein #1 is 1 of 4 human membrane proteins
CC (AAU01210-AAU01213) given in the present invention. These membrane
CC proteins share structural similarity with membrane receptors such as
CC the Ige receptor and mammalian CD20. The novel human membrane proteins
CC are useful for identifying agonists, antagonists and modulators of the
CC membrane proteins, and for producing antibodies specific to the
CC membrane proteins. The membrane proteins can be used for diagnosis,
CC drug screening, pharmacogenomic applications, clinical trial monitoring
CC and the treatment of physiological disorders and diseases. The
CC polynucleotides encoding the membrane proteins can be used to generate
CC PCR primers or probes to identify mutations associated with a particular
CC disease.
XX
SQ Sequence 199 AA;
Query Match 94.2%; Score 726.5; DB 22; Length 199;
Best Local Similarity 74.4%; Pred. No. 3,4e-80;
Matches 148; Conservative 0; Mismatches 0; Indels 51; Gaps 1;
QY 1 MDSSTASPVFLVPPPTITASEVSESTELASTTSTOSPLOKLPARKKIIIGTIOILGIM 60
Db 1 MDSSTASPVFLVPPPTITASEVSESTELASTTSTOSPLOKLPARKKIIIGTIOILGIM 60
QY 61 TFSFGVIFLFTLKPYRPFPIFLSGYFPGSVLFINSGAFLIAVKRKTET----- 112
Db 61 TFSFGVIFLFTLKPYRPFPIFLSGYFPGSVLFINSGAFLIAVKRKTETLILSRIM 120
QY 113 -----LGIITITMTSIIIEFT 129
Db 121 NLSALRAIAGIILLTFEGTILDONYIGYSHONSOKRAVTVLFGIITITMTSIIIEFT 180
QY 130 SLPPSIIIGCHSEDDCCGCC 148
Db 181 SLPPSIIIGCHSEDDCCGCC 199
RESULT 5
AAE10917
ID AAE10917 standard; protein; 201 AA.
XX AAE10917;
AC
XX 18-DEC-2001 (first entry)
DT
XX
XX Human gene 9 encoded immune system-related protein HTENNA5.
DE
XX
XX Human; immune system-related protein; allergy; Rheumatoid arthritis;
KW cancer; Gaucher's disease; viral hepatitis; gene therapy; cytostatic;
KW diabetes mellitus; arrhythmia; wound healing; ischemic lesion; AIDS;
KW Acquired Immune Deficiency Syndrome; viroicide; hepatotropic; vasotropic;

Key	Location/Qualifiers
Domain	21..26
Domain	/label= Immunogenic-epitope
Domain	150..156
Domain	/label= Immunogenic-epitope
MO20016722-A1.	
13-SEP-2001.	
07-MAR-2001; 2001WO-US07260.	
08-MAR-2000; 2000US-187873P.	
11-AUG-2000; 2000US-224367P.	
(HUMA-) HUMAN GENOME SCI INC.	
Ni J, Hilbert D, Kenny JU, Moore PA, Choi GH, Soppet DR, Ebner R; Gruber JR, Endress GA, Ruben SM;	
WPI. 2001-589939/66.	
N-PSDB; AAD18275.	
Novel isolated immune system-related polypeptide useful for treating rheumatoid arthritis, AIDS, allergy, cancer, Gaucher's disease, diabetes mellitus, arrhythmia, wound healing, ischemic lesions and viral hepatitis	
Claim 11; Page 310-311; 315pp; English.	
The invention relates to human immune system-related protein and their DNA. Human immune-system related protein and DNA are useful for preventing, treating or ameliorating a medical condition in a mammalian subject, for diagnosing, preventing or treating immune system-associated disorders, autoimmune disorders (rheumatoid arthritis), inflammatory disorders (allergies), immunodeficiencies (AIDS), inflammatory disorders (allergies), infectious diseases (e.g., viral hepatitis), complement activation disorders, immune complex diseases, neoplastic disorders (cancer), hyperproliferative disorders (Gaucher's disease), disorders associated with neovascularisation, diseases at the cellular level, cardiovascular disorders (arrhythmias), wound healing and epithelial cell proliferation, endocrine disorders (diabetes mellitus) and neurological disorders (ischemic lesions). Immune-system related protein or DNA is useful for preventing hair loss, skin aging due to sunburn, to maintain organs before transplantation, to treat weight disorders, to modulate mammalian characteristics, to change a mammal's mental or physical state, or as a food additive or preservative. Immune-system related DNA is useful in gene therapy, for chromosome identification, radiation hybrid mapping, long range restriction mapping and in forensic biology. The present sequence represents a human immune-system related protein of the invention.	
Sequence 201 AA;	
Query Match	83.7%; Score 645.5; DB 22; Length 201;
Best Local Similarity	72.3%; Pred. No. 2.5e-70;
Matches 136; Conservative	0; Mismatches 1; Indels 51; Gaps 1
1 MDSSAHSPVFLVPPETASSESENEISATFFQSPOKLFARKMKILGTIQLFGIM 60	
1 MDSSAHSPVFLVPPETASSESENEISATFFQSPOKLFARKMKILGTIQLFGIM 60	
TFSEGVILFTLLKRYPRPFIFLISGYPFGWSVLINSAGFLIYARKTKTET----- 112	
TFSEGVILFTLLKRYPRPFIFLISGYPFGWSVLINSAGFLIYARKTKTET----- 112	
61 TFSEGVILFTLLKRYPRPFIFLISGYPFGWSVLINSAGFLIYARKTKTET----- 112	
61 TFSEGVILFTLLKRYPRPFIFLISGYPFGWSVLINSAGFLIYARKTKTET----- 112	
113 -----LGLITLMTFSITLPT 129	

Db	121	NPFLSALGALGIIILFLFPILOQNTICGSHQNSQKAVYVLFGLITLTMTFSLIELFL	180
QY	130	SLPFSILG	137
Db	181	SLPFSIWG	188
RESULT 6			
AB95818			
ID	AB95818 standard; Protein; 212 AA.		
XX			
AC	AB95818;		
XX			
DT	21-JUN-2002 (first entry)		
XX			
DE	Human testicular antigen SEQ ID NO: 1202.		
XX			
KW	Human; testicular antigen; testes; cancer; metastasis; immune disorder		
KW	reproductive system disorder; urinary system disorder; gene therapy;		
KW	cardiovascular disorder; respiratory disorder; neurological disorder;		
KW	gastrointestinal disease; infection; cytostatic.		
XX			
OS	Homo sapiens.		
XX			
PN	W0200155317-A2.		
XX			
PD	02-AUG-2001.		
XX			
PF	17-JAN-2001; 2001WO-US01329.		
XX			
PR	31-JAN-2000;	2000US-0179065.	
PR	04-FEB-2000;	2000US-0180628.	
PR	24-FEB-2000;	2000US-0184664.	
PR	02-MAR-2000;	2000US-0186350.	
PR	16-MAR-2000;	2000US-0189874.	
PR	17-MAR-2000;	2000US-0190076.	
PR	18-APR-2000;	2000US-0198123.	
PR	19-MAY-2000;	2000US-0205515.	
PR	07-JUN-2000;	2000US-0209467.	
PR	28-JUN-2000;	2000US-0214886.	
PR	30-JUL-2000;	2000US-0215135.	
PR	07-JUL-2000;	2000US-0216647.	
PR	07-JUL-2000;	2000US-0216880.	
PR	11-JUL-2000;	2000US-0217487.	
PR	11-JUL-2000;	2000US-0217496.	
PR	14-JUL-2000;	2000US-0218290.	
PR	26-JUL-2000;	2000US-0220963.	
PR	26-JUL-2000;	2000US-0220964.	
PR	14-AUG-2000;	2000US-0224518.	
PR	14-AUG-2000;	2000US-0224519.	
PR	14-AUG-2000;	2000US-0225213.	
PR	14-AUG-2000;	2000US-0225214.	
PR	14-AUG-2000;	2000US-0225266.	
PR	14-AUG-2000;	2000US-0225267.	
PR	14-AUG-2000;	2000US-0225268.	
PR	14-AUG-2000;	2000US-0225270.	
PR	14-AUG-2000;	2000US-0225447.	
PR	14-AUG-2000;	2000US-0225577.	
PR	14-AUG-2000;	2000US-0225758.	
PR	14-AUG-2000;	2000US-0225759.	
PR	18-AUG-2000;	2000US-0226279.	
PR	22-AUG-2000;	2000US-0226681.	
PR	22-AUG-2000;	2000US-0226686.	
PR	22-AUG-2000;	2000US-0227182.	
PR	23-AUG-2000;	2000US-0227709.	
PR	30-AUG-2000;	2000US-0228924.	
PR	01-SEP-2000;	2000US-0229287.	
PR	01-SEP-2000;	2000US-0229343.	
PR	01-SEP-2000;	2000US-0229344.	
PR	01-SEP-2000;	2000US-0229345.	
PR	05-SEP-2000;	2000US-0229509.	
PR	05-SEP-2000;	2000US-0229513.	

XX	Human; reproductive system r
KM	cancer; gene therapy.
XX	
OS	Homo sapiens.
XX	
PN	W0200155320-A2.
XX	
PD	02- <u>AUG</u> -2001.
XX	
PF	17- <u>JAN</u> -2001; 2001WO-US01339.
XX	
PR	31- <u>JAN</u> -2000; 2000US-0178065.
PR	04- <u>FEB</u> -2000; 2000US-0180628.
PR	24- <u>FEB</u> -2000; 2000US-0184664.
PR	02- <u>MAR</u> -2000; 2000US-0186350.
PR	16- <u>MAR</u> -2000; 2000US-0198374.
PR	17- <u>MAR</u> -2000; 2000US-0190076.
PR	18- <u>APR</u> -2000; 2000US-0196123.
PR	19- <u>MAY</u> -2000; 2000US-0205515.
PR	07- <u>JUN</u> -2000; 2000US-0209467.
PR	28- <u>JUN</u> -2000; 2000US-0214886.
PR	30- <u>JUN</u> -2000; 2000US-0215135.
PR	07- <u>JUL</u> -2000; 2000US-0216647.
PR	07- <u>JUL</u> -2000; 2000US-0216880.
PR	11- <u>JUL</u> -2000; 2000US-0217487.
PR	11- <u>JUL</u> -2000; 2000US-0217496.
PR	14- <u>JUL</u> -2000; 2000US-0218290.
PR	26- <u>JUL</u> -2000; 2000US-0220963.
PR	26- <u>JUL</u> -2000; 2000US-0220964.
PR	14- <u>AUG</u> -2000; 2000US-0224518.
PR	14- <u>AUG</u> -2000; 2000US-0224519.
PR	14- <u>AUG</u> -2000; 2000US-0225213.
PR	14- <u>AUG</u> -2000; 2000US-0225214.
PR	14- <u>AUG</u> -2000; 2000US-0225256.
PR	14- <u>AUG</u> -2000; 2000US-0225267.
PR	14- <u>AUG</u> -2000; 2000US-0225268.
PR	14- <u>AUG</u> -2000; 2000US-0225270.
PR	14- <u>AUG</u> -2000; 2000US-0226811.
PR	22- <u>AUG</u> -2000; 2000US-0228182.
PR	22- <u>AUG</u> -2000; 2000US-0228186.
PR	23- <u>AUG</u> -2000; 2000US-0227009.
PR	30- <u>AUG</u> -2000; 2000US-0228924.
PR	01- <u>SEP</u> -2000; 2000US-0229287.
PR	01- <u>SEP</u> -2000; 2000US-0229343.
PR	01- <u>SEP</u> -2000; 2000US-0229344.
PR	01- <u>SEP</u> -2000; 2000US-0229505.
PR	05- <u>SEP</u> -2000; 2000US-0229259.
PR	05- <u>SEP</u> -2000; 2000US-0229513.
PR	06- <u>SEP</u> -2000; 2000US-0230437.
PR	06- <u>SEP</u> -2000; 2000US-0230438.
PR	08- <u>SEP</u> -2000; 2000US-0231242.
PR	08- <u>SEP</u> -2000; 2000US-0231243.
PR	08- <u>SEP</u> -2000; 2000US-0231244.
PR	08- <u>SEP</u> -2000; 2000US-0231413.
PR	08- <u>SEP</u> -2000; 2000US-0231414.
PR	08- <u>SEP</u> -2000; 2000US-0230810.
PR	08- <u>SEP</u> -2000; 2000US-0230811.
PR	12- <u>SEP</u> -2000; 2000US-0231968.
PR	14- <u>SEP</u> -2000; 2000US-0232397.
PR	14- <u>SEP</u> -2000; 2000US-0232398.
PR	14- <u>SEP</u> -2000; 2000US-0233309.
PR	14- <u>SEP</u> -2000; 2000US-0233400.
PR	14- <u>SEP</u> -2000; 2000US-0233401.
PR	14- <u>SEP</u> -2000; 2000US-0233403.
PR	14- <u>SEP</u> -2000; 2000US-0233061.
PR	14- <u>SEP</u> -2000; 2000US-0233064.
PR	14- <u>SEP</u> -2000; 2000US-0233065.
PR	21- <u>SEP</u> -2000; 2000US-0234223.

[illegible]


```

PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX MPI: 2001-465570/50.
XX N-PSDB; AAL01084.
XX
XX Isolated nucleic acid molecule encoding a reproductive system antigen
XX is used in preventing, treating or ameliorating a medical condition
XX
XX Claim 11: SEQ ID NO 3772; 1297pp + Sequence Listing; English.
XX
XX The present invention provides the protein and coding sequences of a
XX number of human reproductive system related antigens. These can be used
XX in the prevention and treatment of reproductive system disorders,
XX including cancer. The present sequence is a protein of the invention.
XX
XX Sequence 212 AA:
XX
XX Query Match 83.7%; Score 645.5; DB 22; Length 212;
XX Best Local Similarity 72.3%; Pred. No. 2.7e-70;
XX Matches 136; Conservative 0; Mismatches 1; Indels 51; Gaps 1;
XX
XX 1 MDSSTAHSPVETVPEPEITASEYESNELSATPTSTQSPLOKLPARKKILGTOLIFGM 60
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||||
XX 12 MDSSTAHSPVETVPEPEITASEYESNELSATPTSTQSPLOKLPARKKILGTOLIFGM 71
XX
XX 61 TFSFGVIFLFTLLKPYRPFPIFLSGYFPMGSLVFLNSGAFILAVRKTTET----- 112
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||||
XX 72 TFSFGVIFLFTLLKPYRPFPIFLSGYFPMGSLVFLNSGAFILAVRKTTETILISRM 131
XX
XX 113 -----LGLILTMFSITELFT 129
XX ||||||||||||||||||
XX
XX 132 NFLSALGAIAGIILNFGFILDQNYICGYSHNSOCKAVTLFLGILITLMTFSITELFT 191
XX
XX 130 SLPEFSILG 137
XX |||||
XX
XX 192 SLPEFSILG 199
XX
XX RESULT 8
XX ABB12234
XX ID ABB12234 standard; peptide; 158 AA.
XX
XX AC ABB12234;
XX
XX DT 11-JAN-2002 (first entry)
XX
XX DE Human secreted protein homologue, SEQ ID NO:2604.
XX
XX Human: cytokine; cell proliferation; cell differentiation; growth factor;
XX haematopoiesis regulation; tissue growth; immunomodulator; activin;
XX inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
XX proliferation; metastasis; cancer; tumour; haematopoietic disorder;
XX myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
XX chronic inflammatory condition; proliferative retinopathy;
XX atherosclerosis; coronary heart disease; arterial ischaemia;
XX bone disorder; osteoporosis; vascular growth disorder;
XX tissue regeneration; wound healing; infection; immune disorder;
XX cell culture; drug screening; gene therapy; anti-inflammatory;
XX antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
XX cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
XX antifungal; vulnerary; antitumor.
XX
XX Homo sapiens.
XX
XX WO200157188-A2.
XX
XX 09-AUG-2001.
XX

```

```

PF 05-FEB-2001; 2001WO-US03800.
XX
XX 03-FEB-2000; 2000US-0496914.
XX 27-APR-2000; 2000US-0560875.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX
XX MPI: 2001-457740/49.
XX N-PSDB; ABA09478.
XX
XX Human proteins and DNA encoding sequences useful for preventing,
XX treating or ameliorating a medical condition in a mammalian subject
XX e.g. arthritis and cancer
XX
XX Claim 20: Page 318; 1963pp; English.
XX
XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
XX sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
XX invention also relates to vectors and recombinant host cells comprising a
XX nucleotide of the invention, methods of producing the novel polypeptides,
XX antibodies against the polypeptides, methods of detecting the nucleotides
XX or polypeptides in a sample, and methods of identifying compounds which
XX bind to polypeptides of the invention. Although novel, many of the
XX polypeptides of the invention have homology to known proteins, thereby
XX giving an insight into their probable biological activities, and hence
XX potential therapeutic applications. The polypeptides of the invention may
XX have various activities, including cytokine, cell proliferation or cell
XX differentiation activities; stem cell growth factor activity;
XX haematopoiesis regulatory activity; tissue growth activity;
XX immunomodulatory activity; activin- or inhibin-related activities;
XX chemotactic or chemokinetic activities; haemostatic, thrombotic or
XX thrombolytic activities; receptor or ligand activities; or may be
XX involved in oncogenesis, cancer cell proliferation or metastasis.
XX depending on their biological activities, polypeptides and nucleotides of
XX the invention are useful for preventing, treating or ameliorating medical
XX conditions, e.g., by protein or gene therapy. Such conditions include
XX cancers, hematopoietic disorders (e.g., myeloid or lymphoid cell
XX disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
XX proliferative retinopathy, atherosclerosis, coronary heart disease,
XX arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
XX vascular growth. Polypeptides involved with tissue regeneration and
XX repair (or nucleic acids encoding them) may be used to promote wound
XX healing (e.g., of burns, incisions and ulcers), while those with
XX immunomodulatory activities may be used in the treatment of viral,
XX bacterial and fungal infections in addition to immune disorders.
XX Polypeptides with growth factor activity may be used in cell cultures to
XX promote cell growth. For example, such polypeptides may be used to
XX manipulate stem cells in culture to give rise to neuroepithelial cells
XX that can be used to augment or replace cells damaged by illness,
XX autoimmune disease or accidental damage. The polypeptides and nucleotides
XX may also be used in the diagnosis of the above conditions, and in drug
XX screening techniques. The present sequence represents a novel human
XX polypeptide of the invention.
XX
XX Sequence 158 AA:
XX
XX Query Match 76.3%; Score 588; DB 22; Length 158;
XX Best Local Similarity 95.9%; Pred. No. 1.9e-63;
XX Matches 117; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
XX
XX 1 MDSSTAHSPVETVPEPEITASEYESNELSATPTSTQSPLOKLPARKKILGTOLIFGM 60
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||||
XX 12 MDSSTAHSPVETVPEPEITASEYESNELSATPTSTQSPLOKLPARKKILGTOLIFGM 71
XX
XX 61 TFSFGVIFLFTLLKPYRPFPIFLSGYFPMGSLVFLNSGAFILAVRKTTETILITM 120
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||||
XX 72 TFSFGVIFLFTLLKPYRPFPIFLSGYFPMGSLVFLNSGAFILAVRKTTETILISRM 131
XX
XX 121 TF 122
XX |
XX
XX 132 NF 133
XX

```

RESULT 9

AAU01212
ID AAU01212 standard; Protein: 138 AA.

XX AC AAU01212;

DT 26-SEP-2001 (first entry)

XX DE Novel human membrane protein #3.

XX KM Human; membrane protein; membrane receptor; IGE receptor; CD20;
XX KW physiological disorder.

XX OS Homo sapiens.

XX PN WO200146417-A2.

XX PD 28-JUN-2001.

XX PF 12-DEC-2000; 2000WO-US33742.

XX PR 22-DEC-1999; 99US-0171567.

XX PA (LEXI-) LEXICON GENETICS INC.

XX PI Walke DW, Turner CA;

XX DR WPI: 2001-408646/43.

XX DR N-PSDB; AAS04281.

PT Polynucleotide encoding novel human membrane protein, useful for
PT identifying agonist, antagonist or modifiers or for producing
PT antibodies useful in therapeutic, diagnostic and pharmacogenomic
PT applications -

XX PS Disclosure; Page 30-31; 32pp; English.

XX CC The present sequence represents novel human membrane protein #3.

XX CC (AAU01210-AAU01213) given in the present invention. These membrane
XX CC proteins share structural similarity with membrane receptors such as
XX CC the IGE receptor and mammalian CD20. The novel human membrane proteins
XX CC are useful for identifying agonists, antagonists and modulators of the
XX CC membrane proteins, and for producing antibodies specific to the
XX CC membrane proteins. The membrane proteins can be used for diagnosis,
XX CC drug screening, pharmacogenomic applications, clinical trial monitoring
XX CC and the treatment of physiological disorders and diseases. The
XX CC polynucleotides encoding the membrane proteins can be used to generate
XX CC PCR primers or probes to identify mutations associated with a particular
XX CC disease.

XX SQ Sequence 138 AA;

Query Match 74.7%; Score 576; DB 22; Length 138;

Best Local Similarity 85.1%; Pred. No. 4.7e-62;

Matches 120; Conservative 2; Mismatches 15; Indels 4; Gaps 1;

QY 1 MOSSTRAHSPVLPPEPTASSESTELSATTFSTOSPLQKLFARKMKLTGTLITLM 60
DB 1 MOSSTRAHSPVLPPEPTASSESTELSATTFSTOSPLQKLFARKMKLTGTLITLM 60

QY 61 TFSFGVIFLTLKPPRPFFLSCGYPWGSVLFNSGAFIAVARKTTELTGLITLM 120
DB 61 TFSFGVIFLTLKPPRPFFLSCGYPWGSVLFNSGAFIAVARKTTELTGLITLM 120

QY 121 TFSIIEFLTSLPFIIGCHSE 141
DB 117 TEVVLITKIVSVRLTSCSME 137

RESULT 10

AAG03074

ID AAG03074 standard; Protein: 77 AA.

XX AC AAG03074;

DT 06-OCT-2000 (first entry)

XX DE Human secreted protein, SEQ ID NO: 7155.

XX KM Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX KW gene therapy; chromosome mapping.

XX OS Homo sapiens.

XX PN EP1033401-A2.

XX PD 06-SEP-2000.

XX PF 21-FEB-2000; 2000EP-0200610.

XX PR 26-FEB-1999; 99US-0122487.

XX PA (GEST) GENSET.

XX PI Dunas Milne Edwards J, Duclert A, Giordano J;

XX DR WPI: 2000-500381/45.

XX DR N-PSDB; AAC03080.

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
PT Claim 13; SEQ ID 7155; 71pp + CD-ROM; English.

XX PS The present sequence is a polypeptide encoded by one of a large number
XX CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
XX CC were prepared from total human RNAs or polyA+ RNAs derived from 30
XX CC different tissues. EST sequences usually correspond mainly to the 3'
XX CC untranslated region (UTR) of the mRNA because they are often obtained
XX CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
XX CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
XX CC those cases where longer cDNA sequences have been obtained, the full 5'
XX CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
XX CC ends and can therefore be used to obtain full length cDNAs and genomic
XX CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
XX CC chromosome mapping procedures. They are used to obtain upstream
XX CC regulatory sequences and to design expression and secretion vectors.

XX SQ Sequence 77 AA;

Query Match 38.1%; Score 294; DB 21; Length 77;

Best Local Similarity 83.6%; Pred. No. 5.2e-26;

Matches 56; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 33 FSTOSPLQKLFARKMKLTGTLITGLIMTFSFGVIFLTLKPPRPFFLSCGYPWGS 92
DB 11 FOLKAPCKKLFARKMKLTGTLITGLIMTFSFGVIFLTLKPPRPFFLSCGYPWGS 70

QY 93 VLFINSG 99
DB 71 VLFINSG 77

RESULT 11

AG89142
ID AAG89142 standard; Protein: 67 AA.

XX AC AAG89142;

XX DT 11-SEP-2001 (first entry)

XX DE Human secreted protein, SEQ ID NO: 262.

DB 61 MILALGV-FLGSLQYRHFQKHFFFTFYTGYPWGVAFCCSGTLVAVAGIKPRTW1Q 119
 QY 113 --LGILITMTFSII-ELFISLPF-----SILGCHSEDDCEQC 148
 DB 120 NSFGMNIASATIALVGTAFSLNIIVNIQSLRCHSSSESPDLC 163

RESULT 15

AAW41056
 ID AAW41056 standard; Protein; 214 AA.

XX AAW41056;

DE 29-APR-1998 (first entry)

XX HTM4 protein.

XX HTM4; antibody; IGE Fc receptor; FcepsilonRIbeta; CD20 antigen; TRAF-1;
 KW 4-transmembrane spanning protein superfamily; ligand binding mimic;
 KW haematopoietic cell detection; inhibitor; tumour necrosis factor; KAP;
 KM TNF receptor-associated factor; CDK-activating kinase; TRAF-2; TRAF-3;
 XX TM4SF.

XX Homo sapiens.

XX US5705615-A.

PD 06-JAN-1998.

XX 03-SEP-1996; 96US-0707340.

XX 03-SEP-1996; 96US-0707340.

XX 06-OCT-1994; 94US-0318492.

XX 03-JUL-1996; 96US-0675648.

XX (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.

XX Adra CN, Lelias J, Lim B;

XX WPI; 1998-086194/08.

XX N-PSDB; AAV03875.

XX Antibody specific for HTM4 protein - useful to block or mimic ligand

XX binding or detect haematopoietic cells

XX Disclosure: Column 25-28; 18pp; English.

XX This sequence is the HTM4 protein. The HTM4 protein is specifically

XX bound by the antibody of the invention. HTM4 shows homology to the beta

XX subunit of the high affinity IGE Fc receptor (FcepsilonRIbeta) and the

XX CD20 antigen, both of which are in the 4-transmembrane spanning proteins

XX superfamily (TM4SF). The antibody of the invention is used to block or

XX mimic binding of ligands such as the tumour necrosis factor (TNF)

XX CC receptor-associated factors TRAF-1, TRAF-2 and TRAF-3 and the

XX CC phosphatase, CDK-activating kinase (KAP), to receptors comprising HTM4,

XX and to detect haematopoietic cells.

XX Sequence 214 AA;

XX

DB 120 NSFGMNIASATIALVGTAFSLNIIVNIQSLRCHSSSESPDLC 163

Search completed: February 24, 2003, 13:02:54
 Job time : 29.6868 secs

Query Match 19.5%; Score 150.5; DB 19; Length 214;
 Best Local Similarity 31.1%; Pred. No. 5.7e-10;
 Matches 51; Conservative 22; Mismatches 64; Indels 27; Gaps 8;

QY 1 MDSSTAH-SPEFLVPPETTASEYSTELSATFTSTQSPLOKLFARKMKILGTLQILFGI 59
 DB 11 LGSASAHGPGSEGEPEELNTSVH-----PINGSPTYK---AKIQVLGAIOILNNA 60

QY 60 MTFSEGVILFLTKPY----PRPFIFLSGYEFGWSVLFINSAGFLIAVKRKTEY---- 112
 DB 61 MILALGV-FLGSLQYRHFQKHFFFTFYTGYPWGVAFCCSGTLVAVAGIKPRTW1Q 119

QY 113 --LGILITMTFSII-ELFISLPF-----SILGCHSEDDCEQC 148

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2003, 13:01:46 ; Search time 17.1264 Seconds
(without alignments)
255.979 Million cell updates/sec

Title: US-09-735-712-8

Perfect score: 771
Sequence: 1 MDSSNAHSPVLVFPPELTA.....SIFPSILGSHEDCDCEQCC 149

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCITUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	150.5	19.5	214	1	US-08-318-492-4
2	150.5	19.5	214	1	US-08-707-340-4
3	150.5	19.5	214	2	US-08-916-902A-3
4	150.5	19.5	214	2	US-08-994-578-4
5	150.5	19.5	214	2	US-09-213-389-3
6	134	17.4	239	2	US-08-916-902A-1
7	134	17.4	239	2	US-09-213-389-1
8	123.5	16.0	243	1	US-07-869-933-29
9	123.5	16.0	243	1	US-07-869-933-33
10	123.5	16.0	243	1	US-08-201-879A-4
11	123.5	16.0	243	2	US-08-916-902A-4
12	123.5	16.0	243	2	US-09-213-389-4
13	123.5	16.0	243	4	US-09-103-663-29
14	123.5	16.0	243	4	US-09-103-663-33
15	123.5	16.0	246	4	US-07-869-933-23
16	123.5	16.0	246	4	US-09-103-663-23
17	123.5	15.8	247	4	US-09-724-864-49
18	117.5	15.2	235	1	US-07-869-933-34
19	117.5	15.2	235	1	US-08-201-879A-5
20	117.5	15.2	235	4	US-09-103-663-34
21	99.5	12.9	244	1	US-07-869-933-32
22	99.5	12.9	244	1	US-08-201-879A-3
23	99.5	12.9	244	4	US-09-103-663-32
24	76	9.9	192	4	US-09-149-476-477
25	75	9.7	325	1	US-08-118-270-30
26	75	9.7	325	5	PCT-US93-08528-30
27	72.5	9.4	360	4	US-08-875-573-20

28	72.5	9.4	360	4	US-09-232-878-2	Sequence 2, Appl
29	72.5	9.4	360	4	US-09-045-583-55	Sequence 55, Appl
30	72.5	9.4	360	4	US-09-534-185-55	Sequence 55, Appl
31	72.5	9.4	438	4	US-08-952-365-2	Sequence 2, Appl
32	69.5	9.0	630	4	US-09-134-001C-4615	Sequence 4615, Ap
33	69	8.9	251	4	US-09-149-476-568	Sequence 568, App
34	69	8.9	423	4	US-09-134-001C-3599	Sequence 3599, Ap
35	68	8.8	86	4	US-09-134-001C-3919	Sequence 3919, Ap
36	68	8.8	272	4	US-09-145-828A-18	Sequence 18, Appl
37	68	8.8	483	1	US-08-194-338-7	Sequence 7, Appl
38	67.5	8.8	275	4	US-09-134-001C-3290	Sequence 3290, Ap
39	67	8.7	618	4	US-08-595-553A-2	Sequence 2, Appl
40	67	8.7	761	4	US-09-625-188-14	Sequence 14, Appl
41	66.5	8.6	351	4	US-09-134-001C-4387	Sequence 4387, Ap
42	65.5	8.5	360	4	US-08-833-752-10	Sequence 10, Appl
43	65.5	8.5	424	4	US-09-149-476-555	Sequence 555, App
44	65.5	8.5	591	1	US-08-484-840-2	Sequence 2, Appl
45	65.5	8.5	591	1	US-08-483-094-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-318-492-4
Sequence 4, Application US/08318492
Patent No. 5352312

GENERAL INFORMATION:
APPLICANT: Llm, Bing

APPLICANT: Adra, Chaker N.

TITLE OF INVENTION: RECOMBINANT HTM4 GENE, PROTEIN AND ASSAYS

TITLE OF INVENTION: ASSAYS

NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESSES:
ADDRESS: Hamilton, Brook, Smith & Reynolds, P.C.

STREET: Two Millitia Drive

CITY: Lexington

STATE: Massachusetts

COUNTRY: U.S.A.

ZIP: 02173

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/318,492

FILING DATE:
CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.

REGISTRATION NUMBER: 22,592

REFERENCE/DOCKET NUMBER: BIH94-03

TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-6240

TELEFAX: (617) 861-9540

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:

LENGTH: 214 amino acids
TYPE: amino acids
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-318-492-4

Query Match 19.5%; Score 150.5; DB 1; Length 214;

Best local similarity 31.1%; Pred. No. 2.8e-10;

Matches 51; Conservative 22; Mismatches 64; Indels 27; Gaps 8;

QY 1 MDSSNAHSPVLVFPPELTAESTELSTQSPQLKIFARKMILGTLQILFGL 59
DB 11 LGSASAHGTPGSEETGPEELNTSYH-----PINGSPOYOK--AKLQVIGAIQLINMA 60

```

QY 60 MTFSEGVFFLTLLKPY----PRPEFLFSGPWPAGSVLFNSAFLIAVKRTTE--- 112
      | : : : : : | : : : : : | : : : : : | : : : : : |
Db 61 MIALAGV-FLSGLQPYHFOKHFFFFFYGYDPTMGAVFRCSSSTLESVAGIKPTRWIQ 119
      | : : : : : | : : : : : | : : : : : | : : : : : |
QY 113 --LGLTILMTFSTI-ELFISLP-----SLTGHSEDCDCEQC 148
      | : : : : : | : : : : : | : : : : : | : : : : : |
Db 120 NSFGNMISATIALGTAFLSLNTAVNIQISLRSCSHSSSESDLC 163

```

RESULT 2
US-08-707-340-4
; Sequence 4, Application US/08707340
; Patent No. 5705615

Query Match	19.5%	Score 150.5	DB 1	Length 214
Best Local Similarity	31.1%	Pred. No. 2.0e-10		
Matches 51; Conservative	22;	Mismatches 64;	Indels 27;	Gaps 8;

RESULT 3
US-08-916-902A-3
; Sequence 3, Application US/08916902A
; Patent No. 5871930

Query Match	19.5%;	Score 150.5;	DB 2;	Length 214;
Best Local Similarity	31.1%;	Pred. No. 2.8e-10;		
Matches 51; Conservative	22;	Mismatches 64;	Indels 27;	Gaps 8;

RESULT 4
US-08-994-578-4
; Sequence 4, Application US/08994578
; Patent No. 5972688
; GENERAL INFORMATION:
; APPLICANT: Lim, Bing


```

APPLICANT: Adra, Chaker N.
APPLICANT: Lelias, Jean-Michel
TITLE OF INVENTION: HTM4 METHODS OF TREATMENT AND ASSAYS.
TITLE OF INVENTION: AGONISTS AND ANTAGONISTS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Milltia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/994,578
FILING DATE: December 19, 1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/707,340
FILING DATE: 03-SEP-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/675,648
FILING DATE: 03-JUL-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/318,492
FILING DATE: 06-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: BIH94-03A2Z
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781) 861-6240
TELEFAX: (781) 861-9540
INFORMATION FOR SEQ ID NO.: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 214 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-994-578-4

Query Match      19.5%; Score 150.5; DB 2; Length 214;
Best Local Similarity 31.1%; Pred. No. 2.8e-10;
Matches 51; Conservative 22; Mismatches 64; Indels 27; Gaps 8;

       1 MDSSNAH-SPEFLVPPPTTASEYESTELSATTFSTQSPLOKLFARKKKITGITIIGFI 59
Db    11 LGSASAHRTPSSENGEPDELNTSVH-----PIGSPDYQR--AKLQVLGAIDILNAA 60
           | :|:|:| | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY   60 MTFSFGVILEFTLKPY---PREPFIFLSGYFWGSVLFINSAGFLIAVKRKTET--- 112
           | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db    61 MILALGV-FLGSLQYPHYFKHFEEFFETGYGPINGAVFCSGFLSVAGIKPRTWIQ 119
QY   113 --LGLILLMTFESII-ELFISLPF-----SIICGSEDCDECQC 148
           ||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db    120 NSFQMNIASATIALVGTAFLSLNLIAVNQTQLRSCHSSSESPPDLC 163

RESULT 5
US-09-213-389-3
Sequence 3, Application US/09213389
Patent No. 5977072
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HIGH AFFINITY IMMUNOGLOBULIN E
TITLE OF INVENTION: RECEPTOR-LIKE PROTEIN
NUMBER OF SEQUENCES: 4
```

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/213,389
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/916,902
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PR-0371 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0535
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 214 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Genbank
CLONE: 561639
US-09-213-389-3

Query Match 19.5%, Score 150.5, DB 2; Length 214;
Best Local Similarity 31.1%; Pred. No. 2.8e-10;
Matches 51; Conservative 22; Mismatches 64; Indels 27; Gaps

QY 1 MDSFAH-SPVFLVPPPTASEYESTELSATFTSTQSPLOKLFARKMKILGTIOIFGI 59
Db 11 LGSAAHSTPGSEGTPEELNYSYH-----PIGSPDYQR---AKLOVLAIOILNAA 60
QY 60 MTFSPGVILFTLLKPY----PRPFPILSGYPFMGSVLFINSGAFLIAVKRRTET--- 112
Db 61 MILAGV-FLSLQGYPHQKHFEEFTFYTGYPINGAVPFCSSGSLSVAGIKPFTWIG 119
QY 113 --LGLITLMTFSIT-ELFISLPF-----SLIGCHSEDCDCQC 148
Db 120 NSFGMNIASATTALVGAFLSLINIAVNTQSLKSSHSSSESPDLG 163

RESULT 6
US-08-916-902A-1
Sequence 1, Application US/08916902A
Patent No. 5871930
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HIGH AFFINITY IMMUNOGLOBULIN E
TITLE OF INVENTION: RECEPTOR-LIKE PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

```

```

:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ for Windows Version 2.0
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/916,902A
: FILING DATE: Herewith
: CLASSIFICATION: 536
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.
: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PF-0371 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-855-0555
: TELEFAX: 415-845-4166
: TELEX:
:
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 239 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: BRAINOT04
: CLONE: 927955
:
: US-08-916-902A-1
:
: Query Match 17.4%; Score 134; DB 2; Length 239;
: Best Local Similarity 31.4%; Pred. No. 3.1e-08;
: Matches 38; Conservative 20; Mismatches 47; Indels 16; Gaps 3;
:
: QY 39 LQKLFAR-KMKILGTIQLFGIMTFSFGVIFLTLKPYPRPFPLSGYPFGSVLEIN 97
:   || : : : || : ||| : : : || : : : ||| : ||
: Db 54 LQKFLKGEPRKVLGYVQILTALMSLSMGITMCMASNTYGSNPISVYIGYITWGSVMFII 113
:   || : || : ||| : : ||| : : ||| : ||
: QY 98 SGAFLIAVRKYTE-----TIGILIT-----LMTFSIIEFLISLPESILCHSED 142
:   || : || : ||| : : ||| : : ||| : ||
: Db 114 SGLSLIAGIRITKGLVNGSLGSMNITSSVLAASGILINTFSLAFSFHHPICNYGNNSN 173
:   || : || : ||| : : ||| : : ||| : ||
: QY 143 C 143
:   |
: Db 174 C 174
:
: RESULT 7
: US-09-213-389-1
: Sequence 1, Application US/09213389
: Patent No. 5977072
: GENERAL INFORMATION:
: APPLICANT: Bandman, Olga
: APPLICANT: Lal, Preeti
: APPLICANT: Corley, Neil C.
: TITLE OF INVENTION: HIGH AFFINITY IMMUNOGLOBULIN E
: TITLE OF INVENTION: RECEPTOR-LIKE PROTEIN
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Inocyte Pharmaceuticals, Inc.
: STREET: 3174 Porter Drive
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/213,389

```

```

: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/916,902
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.
: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PF-0371 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-855-0555
: TELEFAX: 415-845-4166
: TELEX:
:
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 239 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: BRAINOT04
: CLONE: 927955
:
: US-09-213-389-1
:
: Query Match 17.4%; Score 134; DB 2; Length 239;
: Best Local Similarity 31.4%; Pred. No. 3.1e-08;
: Matches 38; Conservative 20; Mismatches 47; Indels 16; Gaps 3;
:
: QY 39 LQKLFAR-KMKILGTIQLFGIMTFSFGVIFLTLKPYPRPFPLSGYPFGSVLEIN 97
:   || : : : || : ||| : : : || : : : ||| : ||
: Db 54 LQKFLKGEPRKVLGYVQILTALMSLSMGITMCMASNTYGSNPISVYIGYITWGSVMFII 113
:   || : || : ||| : : ||| : : ||| : ||
: QY 98 SGAFLIAVRKYTE-----TIGILIT-----LMTFSIIEFLISLPESILCHSED 142
:   || : || : ||| : : ||| : : ||| : ||
: Db 114 SGLSLIAGIRITKGLVNGSLGSMNITSSVLAASGILINTFSLAFSFHHPICNYGNNSN 173
:   || : || : ||| : : ||| : : ||| : ||
: QY 143 C 143
:   |
: Db 174 C 174
:
: RESULT 8
: US-07-869-933-29
: Sequence 29, Application US/07869933
: Patent No. 5770396
: GENERAL INFORMATION:
: APPLICANT: KINER, Jean-Pierre
: TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF
: TITLE OF INVENTION: THE HUMAN B SUBUNIT OF THE HIGH AFFINITY RECEPTOR FOR
: NUMBER OF SEQUENCES: 34
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Foley & Lardner
: STREET: 1800 Diagonal Road, Suite 500
: CITY: Alexandria
: STATE: VA
: COUNTRY: USA
: ZIP: 22313-0299
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/869,933
: FILING DATE: 19920416
: CLASSIFICATION: 514
: ATTORNEY/AGENT INFORMATION:
: NAME: BENT, Stephen A.
: REGISTRATION NUMBER: 29,768
: REFERENCE/DOCKET NUMBER: 40399/154 NIH
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703)836-9300

```


[illegible]

```

ADDRESS: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/213.389
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/916.902
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0371 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 243 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 204117
US-09-213-389-4

Query Match          16.0%; Score 123.5; DB 2; Length 243;
Best Local Similarity 23.0%; Pred. No. 5.8e-07;
Matches    31; Conservative 16; Mismatches 41; Indels 19; Gaps 2;

OY      36 OSPLOKLPARKMKKIIGTQILFGIMTFSGVIFLFTLLKP--YPAPFFFLSGYPPFGSV 93
         |   ::::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |
Db       47 QQTWOSPLKKELKEFGVAGVLGLICLCFGIYCSTLTSDPDDEVLILRAGYFPWCAY 106
         ||::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |
OY      94 LFINSAGFLIAVKRKT-----ETGILITLMFTS 123
         ||::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |
Db      107 LFLVLSGISMSERKNLTLYLVRSLGAINIVSSIAGLGIALIIILNKS 153

RESULT 13
: Sequence 29, Application US/09103663D
: Patent No. 6171803
: GENERAL INFORMATION:
: APPLICANT: Kinet et al.
: TITLE OF INVENTION: Isolation, characterization, and use of the human beta
: TITLE OF INVENTION: subunit of the high affinity receptor for
: TITLE OF INVENTION: immunoglobulin E.
: FILE REFERENCE: 50490
: CURRENT APPLICATION NUMBER: US/09/103.663D
: EARLIER FILING DATE: 1998-06-23
: EARLIER APPLICATION NUMBER: 07/869.933
: NUMBER OF SEQ ID NOS: 35
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 29
: LENGTH: 243
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-103-663-29

```

Best Local Similarity 29.0%; Pred. No. 5.8e-07;
Matches 31; Conservative 16; Mismatches 41; Indels 19; Gaps 2;

QY 36 QSPLOKLFARKKILGTIQLFGIMTFSGVIFLTLKP--YPRPFIFLSGPPWGSV 93

Db 47 QQTWQSFLEKKELEFLGTVLVGLICLFCGTVCSTLQTSDFDEVLLEYRAGYPPWGAV 106

QY 94 LFINSAGFLIAYVRKTT-----ETLGLITLMTFS 123

Db 107 LFVLSGFLISMERRKNTLYVRGSLGANIVSSIAAGLGAIALILINLS 153

RESULT 14

US-09-103-663-33
; Sequence 33, Application US/09103663D
; Patent No. 6171803

; GENERAL INFORMATION:
; APPLICANT: Kinet et al.

; TITLE OF INVENTION: Isolation, characterization, and use of the human beta

; TITLE OF INVENTION: subunit of the high affinity receptor for

; FILE REFERENCE: 50490
; CURRENT APPLICATION NUMBER: US/09/103,663D

; EARLIER FILING DATE: 1998-06-23
; EARLIER APPLICATION NUMBER: 07/869,933

; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 33

; LENGTH: 243

; TYPE: PRT

; ORGANISM: Rattus sp.

US-09-103-663-33

Query Match

Best Local Similarity 16.0%; Score 123.5; DB 4; Length 243;

Matches 31; Conservative 16; Mismatches 41; Indels 19; Gaps 2;

QY 36 QSPLOKLFARKKILGTIQLFGIMTFSGVIFLTLKP--YPRPFIFLSGPPWGSV 93

Db 47 QQTWQSFLEKKELEFLGTVLVGLICLFCGTVCSTLQTSDFDEVLLEYRAGYPPWGAV 106

QY 94 LFINSAGFLIAYVRKTT-----ETLGLITLMTFS 123

Db 107 LFVLSGFLISMERRKNTLYVRGSLGANIVSSIAAGLGAIALILINLS 153

RESULT 15

US-07-869-933-23

; Sequence 23, Application US/07869933

; Patent No. 5770396

; GENERAL INFORMATION:
; APPLICANT: KINET, Jean-Pierre

; TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF

; TITLE OF INVENTION: THE HUMAN B SUBUNIT OF THE HIGH AFFINITY RECEPTOR FOR

; NUMBER OF SEQUENCES: 34

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner

; STREET: 1800 Diagonal Road, Suite 500

; CITY: Alexandria

; STATE: VA

; COUNTRY: USA

; ZIP: 22313-0299

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/869,933

; FILING DATE: 19920416

; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:

; NAME: BENT, Stephen A.

; REGISTRATION NUMBER: 29,768

; REFERENCE/DOCKET NUMBER: 40399/154 NIND

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703)836-9300

; TELEFAX: (703)683-4109

; TELEEX: 899149

; INFORMATION FOR SEQ ID NO: 23:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 246 amino acids

; TYPE: AMINO ACID

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-07-869-933-23

Query Match

Best Local Similarity 16.0%; Score 123.5; DB 1; Length 246;

Matches 31; Conservative 16; Mismatches 41; Indels 19; Gaps 2;

QY 36 QSPLOKLFARKKILGTIQLFGIMTFSGVIFLTLKP--YPRPFIFLSGPPWGSV 93

Db 50 QQTWQSFLEKKELEFLGTVLVGLICLFCGTVCSTLQTSDFDEVLLEYRAGYPPWGAV 109

QY 94 LFINSAGFLIAYVRKTT-----ETLGLITLMTFS 123

Db 110 LFVLSGFLISMERRKNTLYVRGSLGANIVSSIAAGLGAIALILINLS 156

Search completed: February 24, 2003, 13:05:56
Job time : 18.1264 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2003, 13:04:41 : Search time 8.13506 Seconds
(without alignments)
569.068 Million cell updates/sec

Title: US-09-735-712-8

Perfect score: 771

Sequence: 1 MDSSTASHVFLVPPETRA.....SLPFTSLGCHSDCEQCC 149

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 156504 segs, 31069816 residues

Total number of hits satisfying chosen parameters: 156504

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications-AA:*
1: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB pep:*
2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB pep:*
4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB pep:*
5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB pep:*
6: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB pep:*
7: /cgn2_6/ptodata/1/pubppaa/PCT07_PUBCOMB pep:*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB pep:*
9: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB pep:*
10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB pep:*
11: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB pep:*
12: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB pep:*
13: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB pep:*
14: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	771	100.0	149	US-09-735-712-8	Sequence 8, Appl
2	735.5	95.4	200	US-09-924-340-34	Sequence 34, Appl
3	735.5	95.4	200	US-09-992-600A-34	Sequence 34, Appl
4	735.5	95.4	200	US-09-821-821-2	Sequence 2, Appl
5	735.5	95.4	200	US-09-731-872-258	Sequence 258, Appl
6	726.5	94.2	199	US-09-735-712-2	Sequence 2, Appl
7	645.5	83.7	201	US-09-949-842-22	Sequence 22, Appl
8	576	74.7	138	US-09-735-712-6	Sequence 6, Appl
9	263.5	34.2	67	US-09-731-872-262	Sequence 262, Appl
10	221	28.7	81	US-09-735-712-4	Sequence 4, Appl
11	161.3	20.9	167	US-09-765-205-22	Sequence 22, Appl
12	139	18.0	257	US-09-925-302-757	Sequence 757, Appl
13	128	16.6	267	US-09-981-353-82	Sequence 82, Appl
14	128	16.6	299	US-09-739-254-73	Sequence 73, Appl
15	128	16.6	299	US-09-904-615-73	Sequence 73, Appl
16	125	16.2	307	US-09-739-254-142	Sequence 142, Appl
17	125	16.2	307	US-09-904-615-142	Sequence 142, Appl
18	114.5	14.9	250	US-09-736-457-1677	Sequence 1677, Appl
19	114.5	14.9	250	US-09-902-941-1677	Sequence 1677, Appl

20	114.5	14.9	250	9	US-09-849-626-1677	Sequence 1677, Appl
21	114.5	14.9	302	10	US-09-925-297-764	Sequence 764, Appl
22	106.5	13.8	250	9	US-09-902-941-1874	Sequence 1874, Appl
23	106.5	13.8	250	9	US-09-849-626-1874	Sequence 1874, Appl
24	106.5	13.8	286	9	US-09-902-941-1878	Sequence 1878, Appl
25	106.5	13.8	286	9	US-09-849-626-1878	Sequence 1878, Appl
26	106.5	13.8	384	9	US-09-902-941-1876	Sequence 1876, Appl
27	106.5	13.8	384	9	US-09-849-626-1876	Sequence 1876, Appl
28	105.5	13.7	149	9	US-09-796-692-1077	Sequence 1077, Appl
29	105.5	13.7	149	10	US-09-821-821-4	Sequence 4, Appl
30	104.5	13.6	155	9	US-09-796-692-1642	Sequence 1642, Appl
31	99.5	12.9	94	9	US-09-796-692-877	Sequence 877, Appl
32	98.5	12.8	125	9	US-09-796-692-1088	Sequence 1088, Appl
33	98.5	12.8	125	9	US-09-796-692-1562	Sequence 1562, Appl
34	96.5	12.5	204	10	US-09-739-254-117	Sequence 117, Appl
35	96.5	12.5	204	10	US-09-904-615-117	Sequence 117, Appl
36	81	10.5	228	10	US-09-815-242-11542	Sequence 11542, A
37	77	10.0	669	9	US-09-790-852-7	Sequence 7, Appl
38	73.5	9.5	120	9	US-09-796-692-1065	Sequence 1065, Appl
39	72.5	9.4	360	9	US-10-120-394-20	Sequence 20, Appl
40	72.5	9.4	360	9	US-09-764-413-20	Sequence 20, Appl
41	72.5	9.4	360	10	US-09-837-446-2	Sequence 2, Appl
42	72.5	9.4	360	10	US-09-796-744-17	Sequence 17, Appl
43	72.5	9.4	438	10	US-09-894-993-2	Sequence 2, Appl
44	72	9.3	227	1	US-08-834-705-16	Sequence 16, Appl
45	72	9.3	227	10	US-09-815-242-11375	Sequence 11375, A

ALIGNMENTS

```
RESULT 1
US-09-735-712-8
; Sequence 8, Application US/09735712
; Patent No. US20020045743A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; TITLE OF INVENTION: No. US20020045743A1 Human Membrane Proteins and
; FILE REFERENCE: LEX-0109-USA
; CURRENT APPLICATION NUMBER: US/09/735.712
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 60/171,567
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; LENGTH: 149
; SEQ ID NO 8
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-735-712-8

Query Match      100.0%; Score 771; DB 10; Length 149;
Best Local Similarity 100.0%; Pred. No. 1.7e-78;
Matches 149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDSSTASHVFLVPPETRASEYESTELSATFSTQSPLOKLFARKMKILCTIOLFGIM 60
    |||||||
Db 1 MDSSTASHVFLVPPETRASEYESTELSATFSTQSPLOKLFARKMKILCTIOLFGIM 60

QY 61 TFSFGVFLFLKPYRPPIFLSGYPFMGSVFLNSGAFLLVKKRTTTLITIM 120
    |||||||
Db 61 TFSFGVFLFLKPYRPPIFLSGYPFMGSVFLNSGAFLLVKKRTTTLITIM 120

QY 121 TFSITLFLISLPSILCHSDCEQCC 149
    |||||||
Db 121 TFSITLFLISLPSILCHSDCEQCC 149

RESULT 2
US-09-924-340-34
; Sequence 34, Application US/09924340
```

```
; Publication No. US20030027248A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91 US2 REG
; CURRENT APPLICATION NUMBER: US/09/924,340
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: Jpatent
; SEQ ID NO 34
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-924-340-34
```

```
Query Match          95.4%; Score 735.5; DB 9; Length 200;
Best Local Similarity 74.5%; Pred. No. 2.1e-74;
Matches 149; Conservative 0; Mismatches 0; Indels 51; Gaps 1;
```

```
QY 1 MDSTASHPELVFPPEITASEYESTELSTFTSTOSPLOKLPARKMKILGTIQLFGIM 60
    |||||||
DB 1 MDSTASHPELVFPPEITASEYESTELSTFTSTOSPLOKLPARKMKILGTIQLFGIM 60
QY 61 TFSFGVIFLTLLKPYPRPFIFLSGYPFWGSVLFINSGLFLAVRKKTET----- 112
    |||||||
DB 61 TFSFGVIFLTLLKPYPRPFIFLSGYPFWGSVLFINSGLFLAVRKKTETETLILSRIM 120
    |||||||
QY 113 -----LGLITLMTFSITELFT 129
    |||||||
DB 121 NFLSALGAIAGIILLTFEGLDQNYICGYSHONSCKAVTVLFLGILITLMTFSITELFT 180
    |||||||
QY 130 SLPEFSLGCHSEDCDEQCC 149
    |||||||
DB 181 SLPEFSLGCHSEDCDEQCC 200

RESULT 3
US-09-992-600A-34
; Sequence 34, Application US/09992600A
; Publication No. US20030027161A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91 US4 DIV
; CURRENT APPLICATION NUMBER: US/09/992,600A
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/01715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Jpatent
; SEQ ID NO 34
; LENGTH: 200
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
US-09-992-600A-34
```

```
Query Match          95.4%; Score 735.5; DB 9; Length 200;
Best Local Similarity 74.5%; Pred. No. 2.1e-74;
Matches 149; Conservative 0; Mismatches 0; Indels 51; Gaps 1;
```

```
QY 1 MDSTASHPELVFPPEITASEYESTELSTFTSTOSPLOKLPARKMKILGTIQLFGIM 60
    |||||||
DB 1 MDSTASHPELVFPPEITASEYESTELSTFTSTOSPLOKLPARKMKILGTIQLFGIM 60
QY 61 TFSFGVIFLTLLKPYPRPFIFLSGYPFWGSVLFINSGLFLAVRKKTET----- 112
    |||||||
DB 61 TFSFGVIFLTLLKPYPRPFIFLSGYPFWGSVLFINSGLFLAVRKKTETETLILSRIM 120
    |||||||
QY 113 -----LGLITLMTFSITELFT 129
    |||||||
DB 121 NFLSALGAIAGIILLTFEGLDQNYICGYSHONSCKAVTVLFLGILITLMTFSITELFT 180
    |||||||
QY 130 SLPEFSLGCHSEDCDEQCC 149
    |||||||
DB 181 SLPEFSLGCHSEDCDEQCC 200
```

```
RESULT 4
US-09-821-821-2
; Sequence 2, Application US/09821821
; Patent No. US20020064823A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Welch, Andrew A.
; APPLICANT: Calzone, Frank J.
; TITLE OF INVENTION: CD20/19e-Receptor Like Molecules and Uses Thereof
; FILE REFERENCE: 01017/36938A
; CURRENT APPLICATION NUMBER: US/09/821,821
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: US 09/723,258
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: US 60/193,728
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
```

```
; SEQ ID NO 2
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-821-821-2
```

```
Query Match          95.4%; Score 735.5; DB 10; Length 200;
Best Local Similarity 74.5%; Pred. No. 2.1e-74;
Matches 149; Conservative 0; Mismatches 0; Indels 51; Gaps 1;
```

```
QY 1 MDSTASHPELVFPPEITASEYESTELSTFTSTOSPLOKLPARKMKILGTIQLFGIM 60
    |||||||
DB 1 MDSTASHPELVFPPEITASEYESTELSTFTSTOSPLOKLPARKMKILGTIQLFGIM 60
QY 61 TFSFGVIFLTLLKPYPRPFIFLSGYPFWGSVLFINSGLFLAVRKKTET----- 112
    |||||||
DB 61 TFSFGVIFLTLLKPYPRPFIFLSGYPFWGSVLFINSGLFLAVRKKTETETLILSRIM 120
    |||||||
QY 113 -----LGLITLMTFSITELFT 129
    |||||||
DB 121 NFLSALGAIAGIILLTFEGLDQNYICGYSHONSCKAVTVLFLGILITLMTFSITELFT 180
    |||||||
QY 130 SLPEFSLGCHSEDCDEQCC 149
    |||||||
DB 181 SLPEFSLGCHSEDCDEQCC 200

RESULT 5
US-09-731-872-258
; Sequence 258, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
```


? APPLICANT: Dumas Milne Edwards, Jean Baptiste
 ? APPLICANT: Bougueleret, Lydie
 ? APPLICANT: Jobert, Severin
 ? TITLE OF INVENTION: FULL-LENGTH HUMAN CDNA5 ENCODING POTENTIALLY SECRETED PROTEINS
 ? FILE REFERENCE: 78.053.REG
 ? CURRENT APPLICATION NUMBER: US/09/731,872
 ? CURRENT FILING DATE: 2000-12-07
 ? PRIOR APPLICATION NUMBER: US 60/169,629
 ? PRIOR FILING DATE: 1999-12-08
 ? PRIOR APPLICATION NUMBER: US 60/187,470
 ? PRIOR FILING DATE: 2000-03-06
 ? NUMBER OF SEQ ID NOS: 482
 ? SOFTWARE: Patent.pm
 ? SEQ ID NO 258
 ? LENGTH: 200
 ? TYPE: PRT
 ? ORGANISM: Homo sapiens
 ? FEATURE:
 ? NAME/KEY: SIGNAL
 ? LOCATION: 20...1
 ? US-09-731-872-258

Query Match	95.48;	Score 735.5;	DB 10;	Length 200;
Best Local Similarity	74.58;	Pred. No. 2.1e-74;		
Matches 149; Conservative	0;	Mismatches 0;	Indels 51;	Gaps 1;

OY	1	MDSSTASHPVEFLVPPPIITASEYSTSLASTTSTOSTPODLKLRARKKKILGTTIQLFGIM	60
Db	1	MDSSTASHPVEFLVPPPIITASEYSTSLASTTSTOSTPODLKLRARKKKILGTTIQLFGIM	60
OY	61	TFSFGVIFLFTLLKPDPPEPFIFLSGYFPWGSVLFIINSGAFLAVKRKTET-----	112
Db	61	TFSGVIFELFLTKLPYRPRPFIFLSGYFPWGSVLFIINSGAFLAVAKRKTTETILLSRIM	120
OY	113	-----LGIIITYMTFSITLELT	129
Db	121	NFLSALGAIAGIILTFEGFILDQNYICGYSHONSQCKAATVLEGLIITLTMFSITELFI	180
OY	130	SLPFSILGCHEDEDCDCCGCC	149
Db	181	SLPFSILGCHEDEDCDCCGCC	200

```

RESULT 6
US-09-735-712-2
/ Sequence 2, Application US/09735712
/ Patent No. US20020045743A1
/ GENERAL INFORMATION:
/ APPLICANT: Walke, D. Wade
/ APPLICANT: Turner, C. Alexander Jr.
/ TITLE OF INVENTION: No. US20020045743A1el Human Membrane Proteins and
/ TITLE OF INVENTION: Nucleotides Encoding the Same
/ FILE REFERENCE: LEX-0109-USA
/ CURRENT APPLICATION NUMBER: US/09/735,712
/ CURRENT FILING DATE: 2000-12-12
/ PRIOR APPLICATION NUMBER: US 60/171,567
/ PRIOR FILING DATE: 1999-12-22
/ NUMBER OF SEQ ID NOS: 9
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 2
/ LENGTH: 199
/ TYPE: PR1
/ ORGANISM: Homo sapien
/ US-09-735-712-2

```

QY	61	TFSGVIFLEFLLKPYRFRPFIFLSGYPFWGSVLFINSAGFLIAVKRRKET	-----	112
Db	61	TFSGVIFLEFLLKPYRFRPFIFLSGYPFWGSVLFINSAGFLIAVKRRKET		120
QY	113	-----	IGLITLMTFSIIEFLI	129
Db	121	NLSALRAIAGIILLTFEGFIIDQNYICGYSHQNSQCAKAVTLFELGILITLMTFSIIEFLI		160
QY	130	SLPFSIILGCHSEDCDCEQC	148	
Db	181	SLPFSIILGCHSEDCDCEQC	199	

```

RESULT 7
US-09-949-842-22
: Sequence 22, Application US/09949842
: Patent No. US20020164692A1
: GENERAL INFORMATION:
: APPLICANT: N1 et al.
: TITLE OF INVENTION: PT047P1
: FILE REFERENCE: Immune System-Related Polynucleotides, Polypeptides, and Antibodies
: CURRENT APPLICATION NUMBER: US/09/949,842
: CURRENT FILING DATE: 2001-09-02
: PRIOR APPLICATION NUMBER: PCT/US01/07260
: PRIOR FILING DATE: 2001-03-07
: PRIOR APPLICATION NUMBER: 60/224,367
: PRIOR FILING DATE: 2000-08-11
: PRIOR APPLICATION NUMBER: 60/187,873
: PRIOR FILING DATE: 2000-03-08
: NUMBER OF SEQ ID NOS: 26
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 22
: LENGTH: 201
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-949-842-22

```

```

QY      1 MDLTAASPVLVPPPELTASEESTSLKATTTSTOSLPLOKLPARKKKILGTLQILFGIM 60
Db      1 MDSTASPAVLVPPPELTASEESTSLKATTTSTOSLPLOKLPARKKKILGTLQILFGIM 60
QY      61 TFSGCVLFELTLTKPYPPEPIELSGYPFMGSVLEFINSGAFILAVKRKTEET----- 112
Db      61 TFSGCVLFELTLTKPYPPEPIELSGYPFMGSVLEFINSGAFILAVKRKTEETLILLSRIM 120
QY      113 -----LGIILTLMTFSIIELEFI 129
Db      121 NPLSALGALINGIILTFEGFLLDQNYICGYSHQNSQCKAVTVLEFLGILITLMTFSIIELEFI 180
QY      130 SLPSFSLIG 137
Db      181 SLPSFSLIG 188

```

```

: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 2
: LENGTH: 199
: TYPE: PRT
: ORGANISM: Homo sapien
: US-09-735-712-2

Query Match          94.2%; Score 726.5; DB 10; Length 199;
Best Local Similarity 74.4%; Pred. No. 2.1e-73;
Matches 148; Conservative 0; Mismatches 0; Indels 51; Gaps 1;

QY      1 MDSSTASHPVFLVPPPEITASEYESTELSATFESTSTOSPLOKLARKKKIIIGTQILFGIM 60
        |||||||
Db       1 MDSSTASHPVFLVPPPEITASEYESTELSATFESTSTOSPLOKLARKKKIIIGTQILFGIM 60
        |||||||

RESULT 8
US-09-735-712-6
: Sequence 6, Application US/09735712
: Patent No. US20020045743A1
: GENERAL INFORMATION:
: APPLICANT: Walke, D. Wade
: APPLICANT: Turner, C. Alexander Jr.
: TITLE OF INVENTION: No. US20020045743A1el Human Membrane Proteins and
: TITLE OF INVENTION: Polynucleotides Encoding the Same
: FILE REFERENCE: LEX-0109-USA
: CURRENT APPLICATION NUMBER: US/09/735,712
: CURRENT FILING DATE: 2000-12-12
: PRIOR APPLICATION NUMBER: US 60/171,567
: PRIOR FILING DATE: 1999-12-22
: NUMBER OF SEQ ID NOS: 9

```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 138
; TYPE: PRF
; ORGANISM: Homo sapien
US-09-735-712-6
```

```
Query Match
Best Local Similarity 85.1%; Pred. No. 7.1e-57;
Matches 120; Conservative 2; Mismatches 15; Indels 4; Gaps 1;
```

```
QY 1 MDSSTAHSPVFLVPEPEITASEYESTELSTATFTSTQSPLOKLFARKMKILGTIQLIFGIM 60
    |||||
DB 1 MDSSTAHSPVFLVPEPEITASEYESTELSTATFTSTQSPLOKLFARKMKILGTIQLIFGIM 60
QY 61 TFSFGVIFLTLTKPRPRPFILSGYPMGSLVLTNSGAFILAVRKRTETLGLITLM 120
    |||||
DB 61 TFSFGVIFLTLTKPRPRPFILSGYPMGSLVLTNSGAFILAVRKRTETLGLITLM 120
QY 121 TFSIIEFLISLPSILGCHSE 141
    ||:|
DB 117 TFFVILTKIVSVRLLSCSWE 137
```

```
RESULT 9
US-09-731-872-262
; Sequence 262, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bouquelere, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78 US3 REG
; CURRENT APPLICATION NUMBER: US/09/731,872
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 262
; LENGTH: 67
; TYPE: PRF
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -20...-1
US-09-731-872-262
```

```
Query Match
Best Local Similarity 74.7%; Pred. No. 1.6e-22;
Matches 56; Conservative 3; Mismatches 7; Indels 9; Gaps 1;
```

```
QY 1 MDSSTAHSPVFLVPEPEITASEYESTELSTATFTSTQSPLOKLFARKMKILGTIQLIFGIM 60
    |||||
DB 1 MDSSTAHSPVFLVPEPEITASEYESTELSTATFTSTQSPLOKLFARKMKILGTIQLIFGIM 60
```

```
QY 61 TFSFGVIFLTLTKPR 75
    ||:|
DB 55 ---SGALFCSLILPE 66
```

```
RESULT 10
US-09-735-712-4
; Sequence 4, Application US/09735712
; Patent No. US20020045743A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: NO. US20020045743A1a1 Human Membrane Proteins and
; TITLE OF INVENTION: Polynucleotides Encoding the Same
```

```
; FILE REFERENCE: LEX-0109-USA
; CURRENT APPLICATION NUMBER: US/09/735,712
; CURRENT FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 60/171,567
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 81
; TYPE: PRF
; ORGANISM: Homo sapien
US-09-735-712-4
```

```
Query Match
Best Local Similarity 28.7%; Score 221; DB 10; Length 81;
Matches 51; Conservative 8; Mismatches 22; Indels 22; Gaps 3;
```

```
QY 47 MKILGTIQLIFGIMTFSFGVIFLTLTKPRPRPFILSGYPMGSLVLTNSGAFILAVK 106
    ||:|
DB 1 MNLISALRAIAGIILTLFG---FILDQV-----ICGYSHONS-----QCK 38
QY 107 RKTETLGLITLMTFSIIEFLISLPSILGCHSEDCDEQCC 149
    |||||
DB 39 AVTVLEGLITLMTFSIIEFLISLPSILGCHSEDCDEQCC 81
```

```
RESULT 11
US-09-765-205-22
; Sequence 22, Application US/09765205
; Patent No. US20020034800A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Li
; TITLE OF INVENTION: BONE MARROW SECRETED PROTEINS AND POLYNUCLEOTIDES
; FILE REFERENCE: 1458 004/200130 449
; CURRENT APPLICATION NUMBER: US/09/765,205
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US/09/212,440
; PRIOR FILING DATE: 1998-12-16
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 22
; LENGTH: 167
; TYPE: PRF
; ORGANISM: human
US-09-765-205-22
```

```
Query Match
Best Local Similarity 20.9%; Score 161.5; DB 10; Length 167;
Matches 36; Conservative 20; Mismatches 44; Indels 1; Gaps 1;
```

```
QY 39 LQKLFAR-KMKILGTIQLIFGIMTFSFGVIFLTLTKPRPRPFILSGYPMGSLVLTIN 97
    ||:|
DB 35 LQKFLKGPKEKVLGVQILTALMSLSMGITMCMASNTGNSIVYIGYITIMGSVMFII 94
QY 98 SGAFILAVRKRTETLGLITLMTFSIIEFLISLPSILGIC 138
    ||:|
DB 95 SGSLISAAGIRTKGILDGMVLLSVLEFCIAVSLSARGC 135
```

```
RESULT 12
US-09-925-302-757
; Sequence 757, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
```

```
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 757
; LENGTH: 257
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (210)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-302-757

Query Match
Best Local Similarity 26.5%; Pred. No. 5.7e-08;
Matches 41; Conservative 31; Mismatches 73; Indels 10; Gaps 4;

QY 1 MOSSTAHSEVFLVPEPTITASEYESTELSATYFSTQSPLOKLFARKMKILGTIQLFGIM 60
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 26 MTSQVPVNEITIVLPENVI--NFSQAEKEPEPTNQGDSILKHLHAIKIVGTIQLCGMM 83

QY 61 TTSFGYIFLTLKP-YRPFPIFL-SGYFPGSVLFITSGAFLIAVKRKTETL----- 113
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 84 VLSGLIILASASFSFNVGTSTLNSAVPIGPFPIITSGSLSTATEKRLKLVHSSL 143

QY 114 -GILITMTFSIITELFISLPSILGCHSEDCDEQ 147
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 144 VGSILSALSALVGFIITLSYKQATLNPAISIQCELDK 178

RESULT 13
US-09-981-353-82
; Sequence 82, Application US/09981353
; Patent No. US20020160382A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy W.
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
; FILE REFERENCE: PA-0038 US
; CURRENT APPLICATION NUMBER: US/09/981,353
; CURRENT FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 82
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: Incyte ID No. US20020160382A1 611082CD1
US-09-981-353-82

Query Match
Best Local Similarity 16.6%; Score 128; DB 9; Length 267;
Matches 32; Conservative 19; Mismatches 37; Indels 46; Gaps 2;

QY 43 FARKKILGTIQLFGIMFSGVIFL--FTLLKPYRPFPIFLSGYFPGSVLFITSG 99
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 84 FEEENALGVIOIMVGLMHIGFIVCLISFSFREVLGFASTRAVIGGYFPGSLFTISG 143

QY 100 AFLIAVKRKTETL-----GIL 116
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 144 SLSVSASKSLRCIVKSGISGMNIVSSILAFIVGILLVDMCINGVAGQDYMAVLSGKIS 203

QY 117 TILMTFSIITELFIS 130
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 204 ATLMTFSILEFFVA 217

RESULT 14
US-09-739-254-73
; Sequence 73, Application US/09739254
; Patent No. US20010021700A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 49 Human Secreted Proteins
; FILE REFERENCE: P2032P1
; CURRENT APPLICATION NUMBER: US/09/739,254
; EARLIER FILING DATE: 2000-12-19
; EARLIER APPLICATION NUMBER: 09/511,554
; EARLIER FILING DATE: 2000-02-23
; EARLIER APPLICATION NUMBER: PCT/US99/19330
; EARLIER FILING DATE: 1999-08-24
; EARLIER APPLICATION NUMBER: 60/097,917
; EARLIER FILING DATE: 1998-08-25
; EARLIER APPLICATION NUMBER: 60/098,634
; EARLIER FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 73
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (167)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-904-615-73

Query Match
Best Local Similarity 16.6%; Score 128; DB 10; Length 299;
Matches 32; Conservative 19; Mismatches 37; Indels 46; Gaps 2;

QY 43 FARKKILGTIQLFGIMFSGVIFL--FTLLKPYRPFPIFLSGYFPGSVLFITSG 99
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 84 FEEENALGVIOIMVGLMHIGFIVCLISFSFREVLGFASTRAVIGGYFPGSLFTISG 143

QY 100 AFLIAVKRKTETL-----GIL 116
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 144 SLSVSASKSLRCIVKSGISGMNIXSILAFIVGILLVDMCINGVAGQDYMAVLSGKIS 203

QY 117 TILMTFSIITELFIS 130
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 204 ATLMTFSILEFFVA 217

RESULT 15
US-09-904-615-73
; Sequence 73, Application US/09904615
; Patent No. US20020026040A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 49 Human Secreted Proteins
; FILE REFERENCE: P2032P1
; CURRENT APPLICATION NUMBER: US/09/904,615
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/511,554
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/097,917
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: 60/098,634
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 73
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (167)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-904-615-73

Query Match
Best Local Similarity 16.6%; Score 128; DB 10; Length 299;
Matches 32; Conservative 19; Mismatches 37; Indels 46; Gaps 2;
```

```

OY 43 FARKKILCTIOILFCIMTFSGVIFL---FTILKPYRPFIFLISGYPFGSVLEPINS 99
    | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 84 FKBEAKALGVIOIMGLMIGFIVCLISFSFREVIGFASTAVIGYFPWGLSFTISG 143
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 100 AFLIAVKRRKTELE-----GIL 116
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 144 SLVSASAKELSRCLVKSIGMNIKXSIILAFIGVILLVDMCINGVAGODYMAVLSGKIS 203
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 117 TILMTFSITELFIS 130
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 204 ATLMIFSLEFFVA 217
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Search completed: February 24, 2003, 13:11:53
 Job time : 9.13506 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OK protein - protein search, using sw model

Run on: February 24, 2003, 13:01:21 ; Search time 12.4167 Seconds
(without alignments)
1153.613 Million cell updates/sec

Title: US-09-735-712-8

Perfect score: 771
Sequence: 1 MDSSTASPVFLVFPPEITA.....SLPFSILGCHSEDCDCRQCC 149

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	735.5	95.4	200	2 JC7585	testis expressed t
2	150.5	19.5	214	2 IS9258	IgE receptor beta
3	124	16.1	291	2 A30558	B-cell surface ant
4	123.5	16.0	243	2 A31231	high-affinity IgE
5	117.5	15.2	235	2 B34342	IgE Fc receptor be
6	110.5	14.3	297	1 A30586	B-cell surface ant
7	100	13.0	457	2 T37205	hypothetical prote
8	99.5	12.9	244	2 A42806	IgE Fc receptor be
9	92	11.9	455	2 T16070	hypothetical prote
10	88.5	11.5	755	2 T02553	cellulose synthase
11	88	11.4	408	2 C70379	hypothetical prote
12	87	11.3	464	2 T03780	probable integral
13	86.5	11.2	1431	2 T22748	hypothetical prote
14	85.5	11.1	712	2 T02552	cellulose synthase
15	84	10.9	991	2 B71315	conserved hypotet
16	83.5	10.8	555	2 F72555	probable molybden
17	82.5	10.7	448	2 H97008	probable cation ef
18	82.5	10.7	780	2 H84685	probable vacuolar
19	82	10.6	657	2 Q04724	NADH2 dehydrogenas
20	81.3	10.6	590	1 QRB4PR	arginine transport
21	81	10.5	228	2 F71886	hypothetical prote
22	81	10.5	242	2 AB1155	hypothetical prote
23	81	10.5	242	2 AE1513	hypothetical prote
24	80.5	10.4	266	2 T41414	probable receptor-
25	80.5	10.4	662	2 S62707	NADH2 dehydrogenas
26	80	10.4	396	2 AE1796	efflux protein hom
27	80	10.4	458	1 YTBSTK	tetracycline resis
28	80	10.4	458	1 YTBSTK	tetracycline resis
29	80	10.4	458	1 YTSOG	tetracycline resis

30	80	10.4	458	2 JQ1211	tetracycline resis
31	79.5	10.3	244	2 H70193	membrane spanning
32	79.5	10.3	485	2 B98261	L-asparagine peptid
33	79.5	10.3	485	2 AD3023	conserved hypotet
34	79	10.2	292	2 C70421	conserved hypotet
35	79	10.2	555	2 D69770	NADH2 dehydrogenas
36	78.5	10.2	460	2 T13881	protein kinase, pr
37	78.5	10.2	554	2 F90399	probable transport
38	78.5	10.2	573	2 S60912	probable proton pu
39	78.5	10.2	843	2 T06068	conserved hypotet
40	77.5	10.1	144	2 C69348	hypothetical prote
41	77.5	10.1	262	2 S23241	hypothetical prote
42	77	10.0	669	2 T08827	hypothetical prote
43	76.5	9.9	482	2 T17022	NADH2 dehydrogenas
44	76.5	9.9	633	2 T21779	hypothetical prote
45	76	9.9	152	2 JH0751	IgE receptor beta

ALIGNMENTS

```
RESULT 1
JC7585
testis expressed transmembrane-4 protein, TETM4 - human
C:Species: Homo sapiens (man)
C>Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 30-Jun-2001
C:Accession: JC7585
R:Hulet, M.D.; Pagler, E.; Hornby, J.R.; Hogarth, P.M.; Eyre, H.J.; Baker, E.; Craw
Biochem. Biophys. Res. Commun. 280, 374-379, 2001
A:Title: Isolation, tissue distribution, and chromosomal localization of a novel tes
A:Reference number: JC7585; MUID:21092614; PMID:11162526
A:Accession: JC7585
A:Molecule type: mRNA
A:Residues: 1-200 <HUL>
A:Cross-references: GB:AF321127
C:Comment: This protein, a four-transmembrane protein, associates with receptor comp
C:Genetics:
A:Gene: tetm4
A:Map position: 11q12
C:Keywords: signal transduction
F:1-48/Domain: cytoplasmic #status predicted <CYT1>
F:49-70/Domain: transmembrane #status predicted <TM1>
F:71-84/Domain: extracellular #status predicted <EXL1>
F:85-105/Domain: transmembrane #status predicted <TM2>
F:106-118/Domain: intracellular #status predicted <INT>
F:119-138/Domain: transmembrane #status predicted <TM3>
F:139-160/Domain: extracellular #status predicted <EXL2>
F:161-182/Domain: transmembrane #status predicted <TM4>
F:183-200/Domain: cytoplasmic #status predicted <CYT2>

Query Match          95.4% Score 735.5; DB 2; Length 200;
Best local similarity 74.5% Pred. No. 3.3e-67;
Matches 149; Conservative 0; Mismatches 0; Indels 51; Gaps 1;

OY 1 MDSSTASPVFLVFPPEITASEYESTELSATFSTQSPLOLFAKKKILGTIQLIGIM 60
DB 1 MDSSTASPVFLVFPPEITASEYESTELSATFSTQSPLOLFAKKKILGTIQLIGIM 60
OY 1 TFSGVFLFLFLKPYRPFIFLSGYPFMGSVFLINSAGFLINVKKRTET----- 112
DB 1 TFSGVFLFLFLKPYRPFIFLSGYPFMGSVFLINSAGFLINVKKRTET----- 112
OY 61 TFSGVFLFLFLKPYRPFIFLSGYPFMGSVFLINSAGFLINVKKRTET----- 112
DB 61 TFSGVFLFLFLKPYRPFIFLSGYPFMGSVFLINSAGFLINVKKRTET----- 112
OY 113 -----LGLITMTFSIIELEFI 129
DB 121 NLSALGAIAGIILLTFEGFLIDQNYICGYSHONSQCAKRAVVLFLGILITMTFSIIELEFI 180
OY 130 SLPSITLIGCHSEDCDCRQCC 149
DB 181 SLPSITLIGCHSEDCDCRQCC 200

RESULT 2
I59258
```

Ige receptor beta chain / CD20 antigen homolog - human
 C.Species: Homo sapiens (man)
 C.Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 05-Nov-1999
 C.Accession: I59258
 R.Adra, C.N.; Tellas, J.M.; Kobayashi, H.; Kaghad, M.; Morrison, P.; Rowley, J.D.; Lim,
 Proc. Natl. Acad. Sci. U.S.A. 91, 10178-10182, 1994
 A.Title: Cloning of the cDNA for a hematopoietic cell-specific protein related to CD20 a
 ning regions.
 A.Reference number: I59258; MUID:95024008; PMID:7524084
 A.Accession: I59258
 A.Status: Preliminary; translated from GB/EMBL/DBJ
 A.Molecule type: mRNA
 A.Residues: 1-214 <RES>
 A.Cross-references: GB:I5848; NID:g561638; PIDN:AAA62319.1; PID:g561639
 C.Genetics:
 A.Gene: GDB:CD20L; HNM4
 A.Cross-references: GDB:392702
 A.Map position: 11q12-11q13.1
 ;Keywords: immunoglobulin receptor

Query Match	19.5%	Score 150.5;	DB 2;	Length 214;
Best Local Similarity	31.1%	Pred. No. 7.5e+08;		
Matches 51;	Conservative 22;	Mismatches 64;	Indels 27;	Gaps 8;

```

Oy      1 MDSSTAH-SPYFLVFPPEITASEYESTELSATTFSTQSPLOKFLARKMKILGTIQLFGI 59
          : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      11 LGSAAHGTPPSETGPEELNTSYH-----PINSPPDYK--AKLOVLGAIOILNAA 60

```

QY 60 MTFSEGVIFETLLKRY----PREPTEFLSGIPFWGSVLFINSGLIAVKRKTTET--- 112
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 61 MIALGV-FLGSLQYPYHFKRHEFFFEFTYTGYPIWGAVEFCSSGLLSVAGIKPTRWIQ 119

```

OY 113 --LGILITLMTFSII-ELFISLPF-----SILGCHSEDCDCEQC 148
      | : | : : : | : | : | : |
Db 120 NSFGMNIASATIALVGTAFSLNTAVNIQSLRSCHSSSESPPDL 163

```

```

RESULT 3
A30558
B-cell surface antigen CD20 homolog - mouse
N:Alternate names: B-cell differentiation antigen Ly-44
C:Species: Mus musculus (house mouse)
C:Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 07-Feb-1997
C:Accession: A30558
R:Tedder, T.F.; Kleiman, G.; Distche, C.M.; Adler, D.A.; Schlossman, S.F.; Saito, H.
J. Immunol. 141, 4388-4394, 1988
A:Title: Cloning of a complementary DNA encoding a new mouse B lymphocyte differentiation
A:Reference number: A30558; MUID:89067519; PMID:2461992
A:Accession: A30558
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-291 <TED>
C:Genetics:
A:Map position: 19
C:Superfamily: B-cell surface antigen CD20
C:Keywords: membrane protein; phosphoprotein; surface antigen

```

Query Match	16.1%	Score 124	DB 2	Length 291
Best local similarity	33.0%	Pred. NO. 5e-05		
Matches 35	Conservative 18	Mismatches 47	Indels 6	Gaps 3

QY 32 TFSQSPLQKLFARKMKILGTIGLLFGIMSFSGVFLFTLKPYPREFIFLS-CYFWM 900
| | | | | : | | : | : : | : | | | |
Db 27 TSSLVGPTQSFMRRESKALGAVOIMNGLRPHITTLGL---LMIPGVGFAPICLSWVYPLW 822

OY 91 GSVLFINSGAFLIAVKRTETLTGLITLMTFSIIELFSLPSFIL 136
|::|||:::|::|
Dd 83 GGIMYIISGSLAAAEKTSRK-SLVKAKVIMSSLSLFAISGIIL 127

RESULT 4
A31231
high-affinity Ige receptor beta chain - rat

C:Species: *Rattus norvegicus* (Norway rat)
C.Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 05-Nov-1999
C.Accession: A31231
R.Kinetic, J.P.; Blank, U.; Ra, C.; White, K.; Metzger, H.; Kochan, J.
Proc. Natl. Acad. Sci. U.S.A. 85, 6483-6487, 1988
A.Title: Isolation and characterization of cDNAs coding for the beta-subunit of the
A.Reference number: A31231; MUID:88320465; PMID:2970642
A.Accession: A31231
A.Molecule type: mRNA
A.Residues: 1-243 <KIN>
A.Cross-references: GB:M22923; GB:J03845; NID:9204116; PIDN:AAA1149.1; PID:9204117
C:Keywords: Immunoglobulin receptor; transmembrane protein

Query Match	16.0%	Score 123.5	DB 2	Length 243
Best Local Similarity	25.0%	Pred. No. 4.7e-05		
Matches 31	Conservative 16	Mismatches 41	Indels 19	Gaps 2

```

QY      36 QSPLOKLFARKMKILGIIQLLEGMITSFVGIFETLLKP--YPRPFPIFISGYPFWGSV  93
          |      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      47 QQTWQSFLEKLEFLGTVQLVGLICLCFGTVVCSLTQTSDFDDEVLILLRAGYPFWGAV  106

```

```

QY 94 LFINSGLPLAVKRRKTY-----ETLGILITMTFS 123
      ||: || | :|| |
Db 107 LFFVLSGLFSIMSERKNTLYLVRSGLANIVSSIAAGLGAILNLNS 153

```

RESULT 5
B34342

IgE Fc receptor beta chain - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 15-Jun-1990 #sequence_revision 15-Jun-1990 #text_change 05-Nov-1999
 C:Accession: B34342
 R:Ra, C.; Jovan, M.H.E.; Kinet, J.P.
 J. Biol. Chem. 264, 15333-15337, 1989
 A:Title: Complete structure of the mouse mast cell receptor for IgE (Fc-eps
 A:Reference number: A34342; MUID:89359361; PMID:2527850
 A:Accession: B34342
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1235 <RAC>
 A:Cross-references: GB:J05019; NID:9193238; PID:AAA37601.1; PID:g309225
 A:Keywords: immunoglobulin receptor; transmembrane protein

Query Match	15.2%	Score 117.5;	DB 2;	Length 235;
Best Local Similarity	32.8%	Pred No. 0.00018;		
Matches 39;	Conservative 17;	Mismatches 50;		
			Indels 13;	Gaps 6

```

OY 1 MDSSIAHSPVLEVP-DEITASEXESTELSATTFSTQSPLO--KLFAK-MKILGTIQL 566
    ||: | | | : : | | | | | | | | | | | | | | | | | | | | | |
Db 1 MDTEN-RSRADIALPNQOESSADIDELLEASPAKAAPPKQWRTFLKLELEFGATQIL 599

```

QY 57 FGIIMFSEGVLEFLTLKPY-----PREFFILSGPFWGSVLFINSCAFLIAVRKTY 110
:: :: | :| :: |||||::| | |::|
Db 60 VGLICLCFGITIVCSV---YVSDEDEVLLTKLGTPFMGAVLFLVLSGLSIISERKNT 115

RESULT 6
A30586

B-cell surface antigen CD20 - human
N:Alternate names: B-lymphocyte antigen CD20; B1
C:Species: Homo sapiens (man)
C>Date: 08-Jun-1989 #sequence,revision 02-Aug-1996 #text-change 22-Jun-1999
C:Accession: A30586; J10042; A27400; S00387
J:Editor: T.F.; Klejman, G.; Schlossman, S.F.; Salto, H.
J: Immunol. 142, 2560-2568, 1989
A:Title: Structure of the gene encoding the human B lymphocyte differentiation antigen
A:Reference number: A30586; MIM:891762B1; PMID:2466899
A:Accession: A30586
A:Molecule type: DNA
A:Residues: 1-297 <TED>
A:Note: the authors translated the codon ATG for residue 148 as His, TCA for residue
, and GAC for residue 294 as Ser
,; Stamenkovic, I.; Seed, B.

J. Exp. Med. 167, 1975-1980, 1988
 A:Title: Analysis of two cDNA clones encoding the B lymphocyte antigen CD20 (B1.Bp35),
 A:Reference number: J10042; M0ID:8828386; PMID:3260267
 A:Accession: J10042
 A:Molecule type: mRNA
 A:Residues: 1-297 <STA>
 A:Cross-references: GB:12530; NID:q29773; PIDN:CAA31046.1; PID:q29774
 R:Redder, T.F.; Streuli, M.; Schlossman, S.F.; Salto, H.
 Proc. Natl. Acad. Sci. U.S.A. 85, 208-212, 1988
 A:Title: Isolation and structure of a cDNA encoding the B1 (CD20) cell-surface antigen
 A:Reference number: A27400; M0ID:88124792; PMID:2448768
 A:Accession: A27400
 A:Molecule type: mRNA
 A:Residues: 1-297 <TE2>
 A:Cross-references: GB:M27394; GB:J03574; NID:q179307; PIDN:AAA35581.1; PID:q179308
 R:Einfield, D.A.; Brown, J.P.; Valentine, M.A.; Clark, E.A.; Ledbetter, J.A.
 EMBO J. 7, 711-717, 1988
 A:Title: Molecular cloning of the human B cell CD20 receptor predicts a hydrophobic prot
 A:Reference number: S00387; M0ID:88283639; PMID:2456210
 A:Accession: S00387
 A:Molecule type: mRNA
 A:Residues: 1-12, 14-297 <EIN>
 A:Cross-references: EMBL:X07203; NID:g29775; PIDN:CAA30179.1; PID:g29776
 A:Experimental source: Daudi cells
 C:Comment: This protein appears not to be glycosylated. Isoforms of varying molecular we
 C:Comment: This protein plays an important role in B cell activation.
 C:Genetics:
 A:Gene: GDB:CD20
 A:Cross-references: GDB:119761; OMIM:112210
 A:Map position: 11q12-11q13.1
 C:Superfamily: B-cell surface antigen CD20
 C:Keywords: B-cell; phosphoprotein; transmembrane protein
 F:1-51/Domain: intracellular #status predicted <CUT1>
 F:52-103/Domain: intracellular #status predicted <CUT1>
 F:104-116/Domain: transmembrane #status predicted <CUT2>
 F:117-141/Domain: transmembrane #status predicted <CUT2>
 F:142-185/Domain: extracellular #status predicted <EXT3>
 F:186-212/Domain: transmembrane #status predicted <CUT4>
 F:213-297/Domain: intracellular #status predicted <CUT3>
 F:167-183/Disulfide bonds: #status predicted

Query Match 14.3%; Score 110.5; DB 1; Length 297;
 Best Local Similarity 28.8%; Pred. No. 0.0012;
 Matches 30; Conservative 21; Mismatches 46; Indels 7; Gaps 3;

QY 34 STGSPQKLFARKKILGTLQILFGIMTFSGVIFLTLKPPPR-PTIFLSGYPPWGS 92
 DB 36 SLVGPTQSEFFMRKSTLGVQINMGFLHLAGL---LMIPIAGIYAPICVTVWYPLMG 91
 QY 93 VLEFINGAFLIAVRKRTETLILITMTFSIIEFLISLPSFI 136
 DB 92 IMWISGLLAETEKSRKCL-VKCKMIMNSLFLPAISGMIL 133

RESULT 7
 T37205
 hypothetical protein Y5767A.4 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C:Accession: T37205
 R:Ozersky, P.
 submitted to the EMBL Data Library, March 1999
 A:Description: The sequence of C. elegans cosmid Y5767A.
 A:Reference number: Z21634
 A:Accession: T37205
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-457 <ONZ>
 A:Cross-references: EMBL:AF077542; PIDN:AAC26295.1
 C:Genetics:
 A:Map position: II
 A:Introns: 100/2; 193/2; 234/3; 272/2; 313/2; 399/1
 A>Note: Y5767A.4

Query Match 13.0%; Score 100; DB 2; Length 457;
 Best Local Similarity 23.0%; Pred. No. 0.021;
 Matches 41; Conservative 21; Mismatches 36; Indels 80; Gaps 6;

QY 14 FPEPTASEYESTELSATFSTOSPLQKLFARKKILGTLQILFGIMTFSGVIFLTL 73
 DB 197 FPPRYIVSEIMSTARKVSRADOTPLTVT-----ILDGLSKITPAIY----- 240
 QY 74 KPPRPPEFIFLSCYFPGWVLFINSGAFLIAVRKRTETLG-----LITTL 119
 DB 241 --YPIFTF-----LLIQQLRAATALARRKSTSGSRLESTKSDQTKRMVLM 266

QY 120 MTSIIE-----LTSLPFSI-----LGHSE 141
 DB 287 VTFETISGPICIGYILEGTLPPKRSVFRDINDYDLMDAFTIFVAINASVHFLICGVHSO 344

RESULT 8
 A42806
 Ige Fc receptor beta chain - human
 C:Species: Homo sapiens (man)
 C>Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 08-Oct-1999
 C:Accession: A42806; S21154
 R:Kuester, H.; Zhang, L.; Brini, A.T.; MacGlashan, D.W.J.; Kinet, J.P.
 J. Biol. Chem. 267, 12782-12787, 1992
 A:Title: The gene and cDNA for the human high affinity immunoglobulin E receptor beta
 A:Reference number: A42806; M0ID:92316966; PMID:1535625
 A:Accession: A42806
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-244 <KUP>
 A:Cross-references: GB:M89796; NID:g337417; PIDN:AAA60269.1; PID:g337418
 R:Maekawa, K.; Imagawa, N.; Tanaka, Y.; Hatada, S.
 FEBS Lett. 302, 161-165, 1992
 A:Title: Determination of the sequence coding for the beta subunit of the human high-
 A:Reference number: S21154; M0ID:92339505; PMID:1386024
 A:Accession: S21154
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-244 <MAE>
 A:Cross-references: GB:D10583; NID:g219881; PIDN:BA01440.1; PID:d1001914; PID:g21988
 C:Keywords: Immunoglobulin receptor; transmembrane protein
 A:Introns: 19/2; 62/3; 107/3; 126/3; 179/3; 212/3

Query Match 12.9%; Score 99.5; DB 2; Length 244;
 Best Local Similarity 25.8%; Pred. No. 0.013;
 Matches 40; Conservative 20; Mismatches 64; Indels 31; Gaps 6;

QY 9 PVFLVPPETIASSEYESTELSATFSTOSPLQ---KLFARKKILGTLQILFGIMTFSRG 65
 DB 21 PAFEVL--EISPOEVSSGRLLKS--ASSPPLHMTLVLEKDEFLVOTQLTFLMICRG 76
 QY 66 VIFLFLTKPYRPFPT--FLSGYPPGWSVLFINSGAFLIAVRKRTETL----- 113
 DB 77 TVYCVSLDISHIEGDLFFSSFKACYPWGCALFFSISGMLTISRKNATVLYKSGCANTA 136
 QY 114 -----GILITMTFSIIEFLISLPSIILGHSEDC 143
 DB 137 SSISGCTGITI-----LIIINLKSLAVYIHISC 164

RESULT 9
 T16070
 hypothetical protein F14B8.3 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
 C:Accession: T16070
 R:Geisel, C.
 submitted to the EMBL Data Library, June 1995
 A:Description: The sequence of C. elegans cosmid F14B8.
 A:Reference number: Z18456

```

A:Accession: T16070
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-455 <GEL>
A:Cross-references: EMBL:U28737; NID:g960717; PID:g960719; PIDs:AAA68272.1; CESP:F14B8.3
C:Genetics:
A:Introns: 56/3; 193/2; 224/3; 289/2; 302/3; 370/2; 402/3; 442/3

Query Match      11.9%; Score 92; DB 2; Length 455;
Best Local Similarity 29.7%; Pred. No. 0.14;
Matches 27; Conservative 15; Mismatches 29; Indels 20; Gaps 3;

OY    49 ILGTOLIFGIMT-----FSGVLEFTLLKPYPR-PPIFLSGYPFWG  91
       :|::|||::|
Db     303 LLGLVEIYFGILSRNHLISLCINGISMCAAGVILLFGLQKNPTIVLVFLIGY---A 359

OY    92 SVLFINSGAFLIAVKRKTETFTGLITLTMTF 122
       |::|::|::|::|::|::|
Db     360 SIEFIWDACYLCVVEEQVPTEVRGTSCACSF 390

RESULT 10
T02553
cellulose synthase homolog T26B15.10 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 16-Feb-2001
C:Accession: T02553; D84734
R:Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes,
submitted to the EMBL Data Library, July 1998
A:Description: Arabidopsis thaliana chromosome II BAC T26B15 genomic sequence.
A:Reference number: Z14678
A:Accession: T02553
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-755 <ROUT>
A:Cross-references: EMBL:AC004681; NID:g3298532; PID:g3298542
A:Experimental source: cultivar Columbia
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, J.M.; Vanek, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Niemann, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402: 761-768, 1999
A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: D84734
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-755 <STO>
A:Cross-references: GB:AE002093; NID:g3298542; PID:NAC25936.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g32540; T26B15.10
A:Map position: 2
A:Introns: 88/2; 194/3; 236/3; 277/3; 344/3; 390/1; 450/3; 567/3

Query Match      11.5%; Score 88.5; DB 2; Length 755;
Best Local Similarity 30.1%; Pred. No. 0.52;
Matches 41; Conservative 20; Mismatches 44; Indels 31; Gaps 9;

OY    9 PVFLVFPP-----PEIASFEV-STELSATTFSTQSPLOKLFAKKMLILTIOILFGIMTFS 63
       |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db     483 PAFLCGMGOGGEPAVVQORRNATGELLELPFKQSPLIMPCRKIRFRSLAVLY---VFS 539

OY    64 FGVILEFTLLKPYRPPIFLSGYPFWGSVLFINSGAFLIAVKRKTETFTGLITLTIM-- 120
       ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db     540 WG-----LRSELLEFYCLIPAY-----CLIHNSALPFPKV-----YLGIITLVGIN 581

OY    121 -TFSTIELFISPSI 135
       ::::|::|::|::|
Db     582 CLYTIVE-FNNNGFSI 596

```

```

C70379
hypothetical protein aq_917 - Aquifex aeolicus
C:Species: Aquifex aeolicus
C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
C:Accession: C70379
V:Deckert, G.; Warren, F.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;
Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:9819666; PMID:9537320
A:Accession: C70379
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-408 <AAQ>
A:Cross-references: GB:AE000713; NID:g2983424; PIDN:AAC07022.1; PID:g2983440; GB:AE00
C:Genetics:
C:Experimental source: strain VFS
A:Gene: aq_917

Query Match 11.4%; Score 88; DB 2; Length 408;
Best local Similarity 33.3%; Pred. No. 0.31;
Matches 29; Conservative 13; Mismatches 17; Indels 28; Gaps 4;

OY 55 ILFG-----INTFSGVLFELTLKPYRPFITFLSGIPFMGVSVLFINSGAFLAVRKRT 109
      :|::: :|:::|::: :|::: :|:::
Db 185 MLFGTATLVASVSVPSTVFEMFFVTPPYPK-----FISY-----LPKT 222

OY 110 TETLGL-ITLMTFSIITELSLPESI 135
      ||::: :|:::|:::|:::|:::|:::|:::
Db 223 VFTLLIVLKVLPNFSLLIDVFLSLIFTI 249

```

```

RESULT 12
T03780
Probable integral membrane protein - rice
C:/Species: Oryza sativa (rice)
C:/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
C:/Accession: T03780
R/Belouchi, A.; Kwan, T.; Gros, P.
Plant Mol. Biol. 33, 1085-1092, 1997
A:/Title: Cloning and characterization of the OsNrmp family from Oryza sativa, a new
A:/Reference number: Z15079; MUID:97299840; PMID:9154989
A:/Accession: T03780
A:/Status: preliminary; translated from GB/EMBL/DDBJ
A:/Molecule type: mRNA
A:/Residues: 1-464 <BEE>
A:/Cross-references: EMBL:L81152; NID:g2231164; PIDN:AB61961.1; PID:g2231149
A:/Genetics:
C:/Gene: Nrmp2
C:/Superfamily: natural resistance-associated macrophage protein 1

Query Match      11.3% Score 87; DB 2; Length 464;
Best Local Similarity 22.9%; Pred. No. 0.45;
Matches 30; Conservative 22; Mismatches 61; Indels 18; Gaps 2;

QY 14 FPPETASEYSTELSTATFSTQSLOKLFARKKKILGTIOILGIMTFSPGVFLFTLL 73
    :|| ||: : ||| | : ||: ||: ||: ||: ||: ||: ||: ||: ||: ||
Db 81 YPPATALLMTLEALVAGDIEVISAKIKILSAGTVPLMGCVITADDC----- 133

QY 74 KPYPPEPFIISGY-----PFWGSVFINSGLAFILAKRRKTTELTGLITLMFSIIEI 127
    ||||| :||| : ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||
Db 134 -----FILPLENGYGRKLEAFGLVLIAVMVSAFMGEGRPSGKDILLIGLVVPKLSSR 188
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

QY 128 FISLPFSILGC 138
    | :|||
Db 189 TIKAVGIVGC 199

RESULT 13
T22748
Hypothetical protein F55G11.9 - Caenorhabditis elegans
C:/Species: Caenorhabditis elegans
C:/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000

```



```

OY      9  PVELVVRPP-----ELTASEYESTELSATTESTOSPLQKLFARKKKILGTIOILEGIMTFS 63
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      440  PAFIGAFMPGGPEAMLDQRRNATGILEVLFYFKQSPDLGMFCORIKRPROSLAYLY---IFT 496

OY      64  FGVLFLFLKPYPRFPPIFLSGYPFMGSVLFINSGAFLIAVKKRKTETTGILITLM--- 120
      :|  :|  :|  :|  :|  :|  :|  :|  :|  :|  :|  :|  :|  :|  :|  :|  :|
Db      497  WG-----LRSITELIAYCLLPATCLLHNALPFKGYV-----LGIYTVLVMGMH 538

OY      121  -TFSITELFISLPFSI 135
      :|  :|  :|  :|  :|  :|  :|  :|  :|  :|  :|  :|  :|  :|  :|  :|
Db      539  CLYSIME-FMSLGFVS 553

RESULT 15
B71315
conserved hypothetical protein TP0515 - syphilis spirochete
C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 05-Nov-1999
C:Accession: B71315
R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; G-
rison, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; M-
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A:Reference number: A71250; MUID:98332770; PMID:9665876
A:Accession: B71315
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-991 <COL>
A:Cross-references: GB:AE001227; GB:AE000520; NID:93322797; PIDN:AAC65503.1; PID:g332
A:Experimental source: strain Nichols
C:Genetics:
A:Gene: TP0515

Query Match      10.9%; Score 84; DB 2; Length 991;
Best Local Similarity 32.7%; Pred. No. 2;
Matches 36; Conservative 15; Mismatches 45; Indels 14; Gaps 6;

OY      10  VFLVPPPTATSESESTELSATTFST-QSPLOKLFA--RKMKILGTIQLIFGIMTFSFG 65
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      148  VLFQSPPIHSDATSVILKQATLSTEDPEQALMSVHTKNIMWLPPNNQIAFSGVLSFG 207

OY      66  VI-FLFTLLKPYPRFPPIFLSGYPFMGSVLFINSGAFLIAVKKRKTETIG 114
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      208  VVPLLIYDFEYYPKDEFL--GNPVEG--LRSRGAFLI-----QTTTYLLG 248

Search completed: February 24, 2003, 13:05:11
Job time : 14.4167 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2003, 12:59:26 ; Search time 14.1293 Seconds

(without alignments)
437.387 Million cell updates/sec

Title: US-09-735-712-8

Perfect score: 771

Sequence: 1 MDSTHSHSVFLVFPPEIRA.....SLPFTLGCHSDCCDCEGCC 149

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	124	16.1	291	1	CD20_MOUSE
2	123.5	16.0	243	1	FCBE_RAT
3	117.5	15.2	235	1	FCBE_MOUSE
4	110.5	14.3	297	1	CD20_HUMAN
5	99.5	12.9	244	1	FCBE_HUMAN
6	88	11.4	408	1	Y917_AQUAE
7	82	10.6	657	1	NU5M_EMENT
8	81.5	10.6	590	1	CAN1_YEAST
9	80	10.4	458	1	TCR_BACST
10	80	10.4	458	1	TCR_STRAG
11	80	10.4	458	1	TCR_STRPN
12	78.5	10.2	497	1	ALPI_YEAST
13	77.5	10.1	497	1	ANSP_SALTY
14	77	10.0	669	1	COFE_HUMAN
15	75.5	9.8	669	1	NU5M_ARATH
16	75.5	9.8	803	1	ATCO_BACSU
17	75	9.7	601	1	NU5M_MYXGL
18	74.5	9.7	641	1	NU5M_ALIMA
19	74.5	9.7	652	1	NU5M_PODAN
20	74	9.6	402	1	LYCB_BACSU
21	74	9.6	605	1	NU5M_CHICK
22	73.5	9.5	522	1	YOW3_CAEEL
23	73	9.5	439	1	LNT_AQUAE
24	73	9.5	757	1	DRA_MOUSE
25	72.5	9.4	360	1	CKR4_HUMAN
26	72.5	9.4	438	1	MAEL_SCHPO
27	72.5	9.4	576	1	DSBD_PASMU
28	72.5	9.4	1354	1	CYAS_CHICK
29	72	9.3	458	1	TCR_STAHY
30	72	9.3	499	1	ANSP_ECOLI
31	72	9.3	664	1	NU5M_PHYIN
32	71.5	9.3	394	1	EMRD_ECOLI
33	71.5	9.3	459	1	NU4M_MOUSE

34	71	9.2	298	1	NU1M_ARTSF	037714	artemia san
35	71	9.2	315	1	PSIC_HAFIN	P45191	haemophilus
36	71	9.2	373	1	CKR2_RAT	055193	rattus norv
37	71	9.2	362	1	FLAO_METJA	058311	methanococ
38	71	9.2	606	1	NU5M_EQUAS	P92465	equus asinu
39	71	9.2	854	1	VPP2_BOVIN	P97681	bos taurus
40	70.5	9.1	166	1	YCA8_METJA	P81230	methanococ
41	70.5	9.1	349	1	FM12_MACMU	P79191	macaca mula
42	70.5	9.1	459	1	NU4M_PIG	079881	sus scrofa
43	70.5	9.1	527	1	NU2M_ACACA	Q37376	acanthamoeb
44	70	9.1	436	1	SECY_METJA	060175	methanococ
45	70	9.1	604	1	NU5M_HORSE	P48656	equus cabal

ALIGNMENTS

RESULT 1
CD20_MOUSE STANDARD; PRT; 291 AA.
AC P19437;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE B-cell surface protein CD20 homolog (B-cell differentiation antigen
DE LY-44).
GN MS4A1 OR CD20 OR LY-44 OR MS4A2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89067519; PubMed=2461992;
RA Tedder T.F., Klejman G., Distèche C.M., Adler D.A., Schlossman S.F.,
RA Saito H.;
RT "Cloning of a complementary DNA encoding a new mouse B lymphocyte
RT differentiation antigen, homologous to the human B1 (CD20) antigen,
RT and localization of the gene to chromosome 19.";
RL J. Immunol. 141:4388-4394(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX SRRAIN=C57BL/6J; TISSUE=thymus;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Atzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamakawa I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsi G.,
RA Blake J., Boffelli D., Bojtunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Monbaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-Oka K., Wang K.H., Wetz C., Whitaker C., Wilming L.,
RA Yushaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -1- FUNCTION: THIS PROTEIN MAY BE INVOLVED IN THE REGULATION OF
CC B-CELL ACTIVATION AND PROLIFERATION.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- PTM: PHOSPHORYLATED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE MS4A FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

CC      use b non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; M62541; AAA37394.1; -.
DR      DR      AK017903; BAB30996.1; -.
DR      PIR; A30558; A30558.
DR      MGI; 88321; M54a2.
KW      B-cell; Transmembrane; Phosphorylation.
FT      DOMAIN 1 44 CYTOPLASMIC (POTENTIAL).
FT      TRANSMEM 45 65 POTENTIAL.
FT      TRANSMEM 69 89 POTENTIAL.
FT      TRANSMEM 112 132 POTENTIAL.
FT      TRANSMEM 183 203 POTENTIAL.
FT      DOMAIN 204 291 CYTOPLASMIC (POTENTIAL).
SQ      SEQUENCE 291 AA; 31958 MW; DF478BCD2C5C16FC CRC64;

Query Match 16.1%; Score 124; DB 1; Length 291;
Best Local Similarity 33.0%; Pred. NO. 0.00027;
Matches 35; Conservative 18; Mismatches 47; Indels 6; Gaps 3;

QY 32 TSTSTQPLCKLRARKKKILITQILLEGIMTFSEFYIFLTLLKPPRRPFIFLS-GYFPW 90
DB 27 TSLVGPSTQSPFRSKAKGAVQINNGLEFHTLIGSL---LMIPTGVAPICLSWYPLW 82
QY 91 GSVLEFINSGLFIAVKRTTETGLITLMTFSIIEFLFSIPFSL 136
DB 83 GGIMVITISGLLAAAEKTSRK-SLYKAKYIMSSLSLFAISGITL 127

RESULT 2
FCBE_RAT STANDARD; PRT; 243 AA.
ID AC PI3386;
DC 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE High affinity immunoglobulin epsilon receptor beta-subunit (FcER1)
DE (IGF receptor, beta-subunit) (Fc epsilon receptor I beta-chain).
DE MS4A2 OR FCER1B OR FCE1B.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 11-29; 37-43 AND 217-243.
RX MEDLINE=86320465; Pubmed=2970642;
RA Kinet J.-P., Blank U., Ra C., White K., Metzger H., Kochan J.;
RT "Isolation and characterization of cDNAs coding for the beta subunit
RT of the high-affinity receptor for immunoglobulin E.",
RT Proc. Natl. Acad. Sci. U.S.A. 85:6483-6487(1988)
RL -1- FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULIN EPSILON. HIGH
CC AFFINITY RECEPTOR. RESPONSIBLE FOR INITIATING THE ALLERGIC
CC RESPONSE. BINDING OF ALLERGEN TO RECEPTOR-BOUND IGE LEADS TO CELL
CC ACTIVATION AND THE RELEASE OF MEDIATORS (SUCH AS HISTAMINE)
CC RESPONSIBLE FOR THE MANIFESTATIONS OF ALLERGY. THE SAME RECEPTOR
CC ALSO INDUCES THE SECRETION OF IMPORTANT LYMPHOKINES.
CC -1- SUBUNIT: Tetramer of an alpha chain, a beta chain, AND TWO
CC DISULFIDE LINKED GAMMA CHAINS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE MS4A FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; M22923; AAA41149.1; -.
PIR; A31231; A31231.

```

```

KW Ige-binding protein; Receptor; Transmembrane.
FT DOMAIN 1 59 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 60 79 POTENTIAL.
FT DOMAIN 80 97 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 98 117 POTENTIAL.
FT DOMAIN 118 130 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 131 150 POTENTIAL.
FT DOMAIN 151 179 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 180 199 POTENTIAL.
FT DOMAIN 200 243 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 243 AA; 26730 MW; 471DFA59B688E5D CRC64;

Query Match 16.0%; Score 123.5; DB 1; Length 243;
Best Local Similarity 29.0%; Pred. No. 0.00025;
Matches 31; Conservative 16; Mismatches 41; Indels 19; Gaps 2.

OY 36 GSEPLQKARKKKKILGTTILFGIMFSGVYIEFLKLP--YPRPFLSGYPPWGSV 93
Db 47 QQTWQSFLEKLEFLGVTGYLVGLICLGGYVCSLTQTSDEYVLLRYAGYPPWGA 106
OY 94 LEINSGAPLIAVARKTT-----ETGLITLTMFTS 123
Db 107 LEVLGSFLSIMSEKNTLTVLGVSLGANVYSSIAGLTALILNLS 153

RESULT 3
FCER_MOUSE STANDARD; PRT: 235 AA.
AC P20490:
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE High affinity immunoglobulin epsilon receptor beta-subunit (FCER1)
DE (Ige Fc receptor, beta-subunit) (Fc epsilon receptor I beta-chain).
GN MS4A2 OR FCER1B OR FCE1B OR MS4A1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN RP SEQUENCE FROM N.A.
RX MEDLINE=69359361; PubMed=2527850;
RA Ra C., Jouvin M.H.E., Kinet J.-P.;
RT "Complete structure of the mouse mast cell receptor for Ige (Fc
RT epsilon R1) and surface expression of chimeric receptors (rat-mouse-
RL J. Biol. Chem. 264:15323-15327(1989).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=129;
RA Hiraoka S., Watanabe M., Takagaki Y., Fujita-Suzuki K., Shinohara N.,
RA Okumura K., Ra C.;
RT "The genomic structure of the allergy associated Fc receptor beta
RT subunit and its high content of SINS.";
RL submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULIN EPSILON. HIGH
CC AFFINITY RECEPTOR. RESPONSIBLE FOR INITIATING THE ALLERGIC
CC RESPONSE. BINDING OF ALLERGEN TO RECEPTOR-BOUND IGE LEADS TO CELL
CC ACTIVATION AND THE RELEASE OF MEDIATORS (SUCH AS HISTAMINE)
CC RESPONSIBLE FOR THE MANIFESTATIONS OF ALLERGY. THE SAME RECEPTOR
CC ALSO INDUCES THE SECRETION OF IMPORTANT LYMPHOKINES.
CC -1- SUBUNIT: TETRAMER OF AN ALPHA CHAIN, A BETA CHAIN, AND TWO
CC -1- DISULFIDE LINKED GAMMA CHAINS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC -----
DR EMBL: J05019; AAA37601.1; -.
DR EMBL: AB033617; BAA94839.1; -.
DR PIR: B34342; B34342.
DR MGD; MGI:95495; Ms4a1.
KW 1E-binding protein; Receptor; Transmembrane.
FT DOMAIN 1 51 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 52 71 POTENTIAL.
FT DOMAIN 72 89 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 90 109 POTENTIAL.
FT DOMAIN 110 122 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 123 142 POTENTIAL.
FT DOMAIN 143 171 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 172 191 POTENTIAL.
FT DOMAIN 192 235 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 235 AA; 25963 MW; 1C2DBF195738168 CRC64;

Query Match 15.2%; Score 117.5; DB 1; Length 235;
Best Local Similarity 32.8%; Pred. No. 0.00083;
Matches 39; Conservative 17; Mismatches 50; Indels 13; Gaps 6;

QY 1 MDSSTASPVFLVPP-PEITASEFESETELSATFTSTOSPLD--KLPARK-MKILGTQIL 56
DB 1 MDTEN-RSRADIALPNEQESSAPDIELLEASPAKAPPKQWTFPLKLELEFGATQIL 59
QY 57 FGITFSGVIFLFTLKPY-----PRPFIFLSGYPFWGVLINGAFILANKRKT 110
DB 60 VGLICLCEFTIVCSVL--YVSPFDEVLLYKLGIPWGAFLVLSGFLSIISERNKT 115

RESULT 4
CD20_HUMAN STANDARD; PRT; 297 AA.
AC P11836; P08984; Q13963;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE B-lymphocyte antigen CD20 (B-lymphocyte surface antigen B1) (Leu-16)
DE (Bp35).
GN MS4A1 OR CD20.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88258386; PubMed=3260267;
RA Stamenkovic I., Seed B.;
RT "Analysis of two cDNA clones encoding the B lymphocyte antigen CD20 (B1, Bp35), a type III integral membrane protein."
RL J. Exp. Med. 167:1975-1980(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=88124792; PubMed=2448768;
RA Tedder T.F., Klejman M., Schlossman S.F., Saito H.;
RT "Isolation and structure of a cDNA encoding the B1 (CD20) cell-surface antigen of human B lymphocytes."
RL Proc. Natl. Acad. Sci. U.S.A. 85:208-212(1988).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=89176281; PubMed=2466899;
RA Tedder T.F., Klejman G., Schlossman S.F., Saito H.;
RT "Structure of the gene encoding the human B lymphocyte differentiation antigen CD20 (B1).";
RL J. Immunol. 142:2560-2568(1989).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=88283639; PubMed=2456210;
RA Einfeld D.A., Brown J.P., Valentine M.A., Clark E.A., Ledbetter J.A.;
RT "Molecular cloning of the human B cell CD20 receptor predicts a hydrophobic protein with multiple transmembrane domains."
RL Embo J. 7:711-717(1988).
RN [5]

```

```

RP SEQUENCE FROM N.A.
RC TISSUE=Lymph.
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THIS PROTEIN MAY BE INVOLVED IN THE REGULATION OF B-CELL ACTIVATION AND PROLIFERATION.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- PFM: PHOSPHORYLATED. MIGHT BE FUNCTIONALLY REGULATED BY PROTEIN KINASE(S).
CC -1- SIMILARITY: BELONGS TO THE MS4A FAMILY.
CC -1- DATABASE: NAME=PROV; NOTE=CD guide CD20 entry; WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd20.htm".
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb.ch/announce/ or send an email to license@isb.sib.ch).
CC -----
DR EMBL: X12530; CAA31046.1; -.
DR EMBL: M27394; AAA35581.1; -.
DR EMBL: M27395; -. NOT ANNOTATED CDS.
DR EMBL: L23419; AAA88911.1; -.
DR EMBL: L23415; AAA88911.1; JOINED.
DR EMBL: L23416; AAA88911.1; JOINED.
DR EMBL: L23417; AAA88911.1; JOINED.
DR EMBL: X07203; CAA30179.1; -.
DR EMBL: X07204; CAA30180.1; -.
DR EMBL: BC002807; AAH02807.1; -.
DR PIR: A27400; A27400.
DR PIR: J10042; J10042.
DR PIR: A30586; A30586.
DR PIR: S00387; S00387.
DR Genew: HGNC:7315; MS4A1.
DR MIM: 112210; -.
KW B-cell; Transmembrane; Phosphorylation.
FT DOMAIN 1 63 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 64 84 POTENTIAL.
FT TRANSMEM 85 105 POTENTIAL.
FT TRANSMEM 121 141 POTENTIAL.
FT TRANSMEM 189 209 POTENTIAL.
FT DOMAIN 210 297 POTENTIAL.
FT DISULFID 81 167 PROBABLE.
FT DISULFID 167 183 PROBABLE.
FT DISULFID 111 220 PROBABLE.
FT CONFLICT 13 13 P -> L (IN REF. 4).
FT CONFLICT 71 71 M -> I (IN REF. 3).
SQ SEQUENCE 297 AA; 35077 MW; AC5420F8B626BDD1 CRC64;

Query Match 14.3%; Score 110.5; DB 1; Length 297;
Best Local Similarity 28.8%; Pred. No. 0.0045;
Matches 30; Conservative 21; Mismatches 46; Indels 7; Gaps 3;

QY 34 STQSPLOKLARKKKILGTLQILFGIMTFSGVIFLTLKPYRF-PRFPLSGYPWGS 92
DB 36 SLVGPTQSPFMRESKTLGAVQINMGFLPHIALGGL---LMPAIVAPICVTWYPIWGG 91
QY 93 VLFINSGAFILAVKRTETLGLIITLMTFSIELFTSPSIL 136
DB 92 IMYIISGLAATPKNSRKL--VKGMKINSLSLFAISGML 133

RESULT 5
FCER_HUMAN STANDARD; PRT; 244 AA.
ID PCEB_HUMAN
AC 001362;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE High affinity immunoglobulin epsilon receptor beta-subunit (FCER1) (Ige Fc receptor, beta-subunit) (Fc epsilon receptor I beta-chain).

```

GN MS4A2 OR FCER1B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92316966; PubMed=1535625;
 RA Kuester H., Zhang L., Brini A.T., Macglashan D.W., Kinet J.-P.;
 RT "The gene and cDNA for the human high affinity immunoglobulin E
 RT receptor beta chain and expression of the complete human receptor.";
 RL J. Biol. Chem. 267:12782-12787(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92316966; PubMed=1386024;
 RA Maekawa K., Imagawa N., Tanaka Y., Harada S.;
 RT "Determination of the sequence coding for the beta subunit of the
 RT human high-affinity IgE receptor.";
 RL FEBS Lett. 302:161-165(1992).
 RN [3]
 RP VARIANT GLY-237.
 RX MEDLINE=96414302; PubMed=8817330;
 RA Hill M.R., Cookson W.O.;
 RT "A new variant of the beta subunit of the high-affinity receptor for
 RT immunoglobulin E (Fc epsilon RI-beta E2376): associations with
 RT measures of atopy and bronchial hyper-responsiveness.";
 RL Hum. Mol. Genet. 5:959-962(1996).
 RN [4]
 RP VARIANT GLY-237.
 RX MEDLINE=96440420; PubMed=8842731;
 RA Shirakawa T., Mao X.-Q., Sasaki S., Enomoto T., Kawai M., Morimoto K.,
 RA Hopkin J.;
 RT "Association between atopic asthma and a coding variant of Fc-epsilon-
 RT RI-beta in a Japanese population.";
 RL Hum. Mol. Genet. 5:1129-1130(1996).
 RN [5]
 RP ERRATUM.
 RX MEDLINE=97123518; PubMed=8968765;
 RA Shirakawa T., Mao X.-Q., Sasaki S., Enomoto T., Kawai M., Morimoto K.,
 RA Hopkin J.;
 RT Hum. Mol. Genet. 5:2068-2068(1996).
 CC -1- FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULIN Epsilon. HIGH
 CC AFFINITY RECEPTOR. RESPONSIBLE FOR INITIATING THE ALLERGIC
 CC RESPONSE. BINDING OF ALLERGEN TO RECEPTOR-BOUND IGE LEADS TO CELL
 CC ACTIVATION AND THE RELEASE OF MEDIATORS (SUCH AS HISTAMINE)
 CC RESPONSIBLE FOR THE MANIFESTATIONS OF ALLERGY. THE SAME RECEPTOR
 CC ALSO INDUCES THE SECRETION OF IMPORTANT LYMPHOKINES.
 CC -1- SUBUNIT: TETRAMER OF AN ALPHA CHAIN, A BETA CHAIN, AND TWO
 CC DISULFIDE LINKED GAMMA CHAINS.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: FOUND ON THE SURFACE OF MAST CELLS AND
 CC BASOPHILS.
 CC -1- POLYMORPHISM: VARIANT GLU-237 HAS BEEN FOUND TO BE RESENT IN ABOUT
 CC 5.3% OF A 1004 INDIVIDUALS POPULATION SAMPLE IN AUSTRALIA. IT
 CC SEEMS TO BE A RISK FACTOR FOR ATOPIC DERMATITIS AND ASTHMA.
 CC -1- SIMILARITY: BELONGS TO THE MS4A FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: D10583; BAA01440.1; -
 DR EMBL: M89796; AAA60269.1; -
 DR PIR: S21134; S21134.
 DR PIR: A42806; A42806.
 DR Genew: HGNC:7316; MS4A2.
 DR MIM: 147138; -
 KM IGE-binding protein; Receptor; Transmembrane; Polymorphism.
 FT DOMAIN 1 59 CYTOPLASMIC (POTENTIAL).
 FT

FT TRANSMEM 60 79 POTENTIAL.
 FT DOMAIN 80 97 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 98 117 POTENTIAL.
 FT DOMAIN 118 130 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 131 150 POTENTIAL.
 FT DOMAIN 151 180 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 181 200 POTENTIAL.
 FT DOMAIN 201 244 CYTOPLASMIC (POTENTIAL).
 FT VARIANT 237 237 E -> G.
 FT
 SQ SEQUENCE 244 AA; 26533 MW; CE523102D5F567AF CRC64;
 Query Match 12.9%; Score 99.5; DB 1; Length 244;
 Best Local Similarity 25.8%; Pred. No. 0.036;
 Matches 40; Conservative 20; Mismatches 64; Indels 31; Gaps 6;
 QY 9 PVFLVPPPTASEYSTELATFTSQPLQ---KLPAKMKILGTIILGIMFSPRG 65
 DB 21 PAFEVL--EISPOVSSGRILKS--ASSPLHTWLTVLKQDFELVLTITLMTICLFG 76
 QY 66 VIFLTLKPYRPFPT--FLSGYPGWSYLFINSGLFIYAVKRTETL----- 113
 DB 77 TVCVSLVDISHIGDIFSSPKACYPFGALFFISGMLSTISRRAATYLVKSGLCANRA 136
 QY 114 ----GILITLMTFTILEFLISLPSILICHSKEDC 143
 DB 137 SSISGIGTIT-----LITLKKSLAYIHHSK 164
 RESULT 6
 ID Y917_AQUAE STANDARD; PRT; 408 AA.
 AC 067062;
 DT 16-OCT-2001 (Rel. 40; Created)
 DT 16-OCT-2001 (Rel. 40; Last sequence update)
 DT 16-OCT-2001 (Rel. 40; Last annotation update)
 DE Hypothetical protein AQ_917.
 GN AQ_917.
 OS Aquifex aeolicus.
 OC Bacteria; Aquificae; Aquificae (class); Aquificales; Aquificaceae;
 OC Aquifex.
 CC NCBI_TaxID=63363;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-VF5;
 RX MEDLINE=98196666; PubMed=9537320;
 RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
 RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Anjay M., Huber R.,
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
 RT "The complete genome of the hyperthermophilic bacterium Aquifex
 RT aeolicus.";
 RL Nature 392:353-358(1998).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AE000713; AAC07022.1; -
 KW Hypothetical protein; Transmembrane; Complete proteome.
 FT TRANSMEM 15 37 POTENTIAL.
 FT TRANSMEM 50 72 POTENTIAL.
 FT TRANSMEM 87 104 POTENTIAL.
 FT TRANSMEM 111 130 POTENTIAL.
 FT TRANSMEM 145 167 POTENTIAL.
 FT TRANSMEM 180 202 POTENTIAL.
 FT TRANSMEM 236 258 POTENTIAL.
 FT TRANSMEM 270 289 POTENTIAL.
 FT TRANSMEM 304 326 POTENTIAL.
 FT

[illegible]

Db 268 EYSSVTVLVCIMLGATITTVFSSLIQLFQOILKKLFPAVSTMQGMVIAVGLSYN---LA 324

Qy 69 LFTLLKPRPRPFPIFLSISYFPMGSLVLFNNSCAFILAVR-----KR---TTEFLGLITLIM 120

Db 325 LFLHV-----NHAFFYKALLFLGASVTHAVADNODFRKRYGGLREFPLITYVM 372

Qy 121 TFSITFLFSLPSFLSGHSED 142

Db 373 LIASLST-VAVPF-WTGFFYSKD 392

RESULT 8

ID	CAN1_YEAST	STANDARD:	PRT:	590 AA.
AC	P04817:			
DT	13-AUG-1987 (Rel. 05, Created)			
DT	01-FEB-1995 (Rel. 31, Last sequence update)			
DT	15-JUL-1998 (Rel. 36, Last annotation update)			
DE	Arginine permease.			
GN	CAN1 OR YEL063C.			
OS	Saccharomyces cerevisiae (Baker's yeast).			
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;			
OC	Saccharomycetales; Saccharomycetaceae; Saccharomycos.			
OX	NCBI_TaxID=4932;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88165106; PubMed=3327612;			
RA	Ahmad M., Bussey H.:			
RT	"Yeast arginine permease: nucleotide sequence of the CAN1 gene.";			
RL	Curr. Genet. 10:587-592(1986).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=66008235; PubMed=3900064;			
RA	Hoffmann W.:			
RT	"Molecular characterization of the CAN1 locus in Saccharomyces			
RT	cerevisiae. A transmembrane protein without N-terminal hydrophobic			
RL	signal sequence.";			
RL	J. Biol. Chem. 260:11831-11837(1985).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ES288C / AB972;			
RA	Dietsch F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,			
RA	Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,			
RA	Chung E., Duncan M., Guzman E., Hartzell G., Hunnicke-Smith S.,			
RA	H्यान R., Kayser A., Komp C., Lahekari D., Lew H., Lin D.,			
RA	Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,			
RA	Petel F.X., Roberts D., Sehl P., Schramm S., Shogren R., Smith V.,			
RA	Taylor P., Wei Y., Yellon M., Bostein D., Davis R.W.:			
RL	Submitted (DEC-1994) to the EMBL/Genbank/DBJ databases.			
CC	-1- FUNCTION: HIGH-AFFINITY PERMEASE FOR ARGinine.			
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.			
CC	-1- SIMILARITY: BELONGS TO THE AMINO ACID PERMEASE FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; X03784; CAA27416.1; -;			
DR	EMBL; M11724; AAA34467.1; -;			
DR	EMBL; U18795; AAB65024.1; -;			
DR	PIR; A23922; QRYPR.			
DR	SGD; S0000789; CAN1.			
DR	InterPro: IPR002293; AA/permease1.			
DR	InterPro: IPR004840; AAC permease.			
DR	InterPro: IPR004841; permease.			
DR	InterPro: IPR004762; Yeast_AA_perm.			
DR	Pfam; Pf00324; aa_permeases; 1.			
DR	TIGRFAMS; TIGR00913; ZAO310; 1.			
DR	PROSITE; PS00218; AMINO_ACID_PERMEASE_1; 1.			

KW Transport; Amino-acid transport; Transmembrane.
 FT DOMAIN 1 92 HYDROPHILIC.
 FT TRANSMEM 93 110 POTENTIAL.
 FT TRANSMEM 230 254 POTENTIAL.
 FT TRANSMEM 324 343 POTENTIAL.
 FT TRANSMEM 377 396 POTENTIAL.
 FT TRANSMEM 420 437 POTENTIAL.
 FT TRANSMEM 451 469 POTENTIAL.
 FT TRANSMEM 494 518 POTENTIAL.
 FT TRANSMEM 525 548 POTENTIAL.
 FT DOMAIN 549 590 HYDROPHILIC.
 FT CONFLICT 534 534 I->V (IN REF. 1).
 SQ SEQUENCE 590 AA; 65785 MW; 4E5A21C77145330D CRC64;

Query Match 10.6%; Score 81.5; DB 1; Length 590;
 Best Local Similarity 22.9%; Pred. No. 3.5;
 Matches 27; Conservative 26; Mismatches 42; Indels 23; Gaps 4;

OY 20 ASEYSELSATSTQSPLOKLFARKKILGTIQLGIMTFSGVIFLTLKPY--P 77
 DB 294 AFTFOGTELVGITEAGEANPRKSVRAIK-----KVFRILTFYIGSLIFGLVYPNDP 348
 OY 78 RPPFLFSGYPPWGSVFLINSAGFLIAVKRTTEFLG-----LITMTFSIIEFLI 129
 DB 349 KLT-----QSTSVSTSPFLIALENSGTQVLPHEINAVLITTIISANSNIVY 396

RESULT 9

TCR_BACST
 ID TCR_BACST STANDARD: PRT; 458 AA.

AC P07561;

DT 01-APR-1988 (Rel. 07, Created)

DT 01-APR-1988 (Rel. 07, Last sequence update)

DT 01-OCT-1994 (Rel. 30, Last annotation update)

DE Tetracycline resistance protein.

GN TET.

OS Bacillus stearothermophilus.

OG Plasmid pTH15.

OC Bacteria; Firmicutes; Bacillales; Geobacillus.

OX NCBI_TaxID=1422;

RN [1]

RP SEQUENCE FROM N.A. Pubmed=2996983;

RX Hoshino T., Ikeda T., Tomizuka N., Furukawa K.;

RT "Nucleotide sequence of the tetracycline resistance gene of pTH15, a thermophilic Bacillus plasmid: comparison with staphylococcal TCR controls";

RL Gene 37:131-138(1985).

CC -I- FUNCTION: RESISTANCE TO TETRACYCLINE BY AN ACTIVE TETRACYCLINE EFFLUX. THIS IS AN ENERGY-DEPENDENT PROCESS THAT DECREASES THE ACCUMULATION OF THE ANTIBIOTIC IN WHOLE CELLS. THIS PROTEIN FUNCTIONS AS A METAL-TETRACYCLINE/H+ ANTIEXPORTER.

CC -I- SUBCELLULAR LOCATION: Integral membrane protein.

CC -I- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN AS THE DRUG RESISTANCE TRANSDUCASE FAMILY).

CC -I- SUBCELLULAR LOCATION: Integral membrane protein.

CC -I- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN AS THE DRUG RESISTANCE TRANSDUCASE FAMILY).

CC -I- SUBCELLULAR LOCATION: Integral membrane protein.

CC -I- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN AS THE DRUG RESISTANCE TRANSDUCASE FAMILY).

CC -I- SUBCELLULAR LOCATION: Integral membrane protein.

CC -I- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN AS THE DRUG RESISTANCE TRANSDUCASE FAMILY).

CC -I- SUBCELLULAR LOCATION: Integral membrane protein.

CC -I- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN AS THE DRUG RESISTANCE TRANSDUCASE FAMILY).

CC -I- SUBCELLULAR LOCATION: Integral membrane protein.

CC -I- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN AS THE DRUG RESISTANCE TRANSDUCASE FAMILY).

CC -I- SUBCELLULAR LOCATION: Integral membrane protein.

CC -I- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN AS THE DRUG RESISTANCE TRANSDUCASE FAMILY).

CC -I- SUBCELLULAR LOCATION: Integral membrane protein.

CC -I- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN AS THE DRUG RESISTANCE TRANSDUCASE FAMILY).

CC -I- SUBCELLULAR LOCATION: Integral membrane protein.

CC -I- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN AS THE DRUG RESISTANCE TRANSDUCASE FAMILY).

CC -I- SUBCELLULAR LOCATION: Integral membrane protein.

CC -I- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN AS THE DRUG RESISTANCE TRANSDUCASE FAMILY).

CC -I- SUBCELLULAR LOCATION: Integral membrane protein.

CC -I- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN AS THE DRUG RESISTANCE TRANSDUCASE FAMILY).

CC -I- SUBCELLULAR LOCATION: Integral membrane protein.

CC -I- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN AS THE DRUG RESISTANCE TRANSDUCASE FAMILY).

CC -I- SUBCELLULAR LOCATION: Integral membrane protein.

CC -I- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN AS THE DRUG RESISTANCE TRANSDUCASE FAMILY).

FT TRANSMEM 165 185 POTENTIAL.
 FT TRANSMEM 201 221 POTENTIAL.
 FT TRANSMEM 223 240 POTENTIAL.
 FT TRANSMEM 256 276 POTENTIAL.
 FT TRANSMEM 297 317 POTENTIAL.
 FT TRANSMEM 324 344 POTENTIAL.
 FT TRANSMEM 346 365 POTENTIAL.
 FT TRANSMEM 432 451 POTENTIAL.
 SQ SEQUENCE 458 AA; 50119 MW; E1F33F4C57E7B06A CRC64;

Query Match 10.4%; Score 80; DB 1; Length 458;
 Best Local Similarity 25.8%; Pred. No. 3.8;
 Matches 32; Conservative 21; Mismatches 41; Indels 30; Gaps 6;

OY 32 TFSQSPLOKLFARKKILGTIQLGIMTFSGVIFLTLKPYRPPFLFSGYPPWG 91
 DB 178 TITTYPFLMKLLKRVKRGKHFDIR-GIILMSVGIVF-----MLFTTSYST-- 223
 OY 92 SVLFINSAGFLIAVK--RKTE-----TLGILITLMTFSIIEFLISL-PSI 135
 DB 224 SFLIYVLSFLIFVKHKKVTDPEVDPLGKNIPMIGVLCGLIFGVAGFVSNVPM 283
 OY 136 LGCH 139
 DB 284 KDVA 287

RESULT 10

TCR_STRAG
 ID TCR_STRAG STANDARD: PRT; 458 AA.

AC P13924;

DT 01-JAN-1990 (Rel. 13, Created)

DT 01-JAN-1990 (Rel. 13, Last sequence update)

DT 01-OCT-1994 (Rel. 30, Last annotation update)

DE Tetracycline resistance protein.

GN TET.

OS Streptococcus agalactiae.

OG Plasmid pMV158.

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OC Streptococcus.

OX NCBI_TaxID=1311;

RN [1]

RP SEQUENCE FROM N.A. Pubmed=2677995;

RX MEDLINE=90016790; Pubmed=2677995;

RT "Similarity of plasmids pUB10, pTB913 and pMV158";

RL Nucleic Acids Res. 17:7283-7294(1989).

CC -I- FUNCTION: RESISTANCE TO TETRACYCLINE BY AN ACTIVE TETRACYCLINE EFFLUX. THIS IS AN ENERGY-DEPENDENT PROCESS THAT DECREASES THE ACCUMULATION OF THE ANTIBIOTIC IN WHOLE CELLS. THIS PROTEIN FUNCTIONS AS A METAL-TETRACYCLINE/H+ ANTIEXPORTER.

CC -I- SUBCELLULAR LOCATION: Integral membrane protein.

CC -I- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN AS THE DRUG RESISTANCE TRANSDUCASE FAMILY).

CC -I- SUBCELLULAR LOCATION: Integral membrane protein.

CC -I- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN AS THE DRUG RESISTANCE TRANSDUCASE FAMILY).

CC -I- SUBCELLULAR LOCATION: Integral membrane protein.

CC -I- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN AS THE DRUG RESISTANCE TRANSDUCASE FAMILY).

CC -I- SUBCELLULAR LOCATION: Integral membrane protein.

CC -I- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN AS THE DRUG RESISTANCE TRANSDUCASE FAMILY).

CC -I- SUBCELLULAR LOCATION: Integral membrane protein.

CC -I- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN AS THE DRUG RESISTANCE TRANSDUCASE FAMILY).

CC -I- SUBCELLULAR LOCATION: Integral membrane protein.

CC -I- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN AS THE DRUG RESISTANCE TRANSDUCASE FAMILY).

CC -I- SUBCELLULAR LOCATION: Integral membrane protein.

CC -I- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN AS THE DRUG RESISTANCE TRANSDUCASE FAMILY).

CC -I- SUBCELLULAR LOCATION: Integral membrane protein.

CC -I- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN AS THE DRUG RESISTANCE TRANSDUCASE FAMILY).

CC -I- SUBCELLULAR LOCATION: Integral membrane protein.

CC -I- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN AS THE DRUG RESISTANCE TRANSDUCASE FAMILY).

CC -I- SUBCELLULAR LOCATION: Integral membrane protein.

CC -I- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN AS THE DRUG RESISTANCE TRANSDUCASE FAMILY).

CC -I- SUBCELLULAR LOCATION: Integral membrane protein.

CC -I- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN AS THE DRUG RESISTANCE TRANSDUCASE FAMILY).

CC -I- SUBCELLULAR LOCATION: Integral membrane protein.

CC -I- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN AS THE DRUG RESISTANCE TRANSDUCASE FAMILY).

CC -I- SUBCELLULAR LOCATION: Integral membrane protein.

FT	TRANSMEM	165	185	POTENTIAL.
FT	TRANSMEM	201	221	POTENTIAL.
FT	TRANSMEM	223	240	POTENTIAL.
FT	TRANSMEM	256	276	POTENTIAL.
FT	TRANSMEM	297	317	POTENTIAL.
FT	TRANSMEM	324	344	POTENTIAL.
FT	TRANSMEM	346	365	POTENTIAL.
FT	TRANSMEM	432	451	POTENTIAL.
SO	SEQUENCE	458 AA;	50006 MM;	A02014E7CA19995 CRC64;

Query Match 10.4%; Score 80; DB 1; Length 458;
Best Local Similarity 25.8%; Pred. NO. 3.8;
Matches 32; Conservative 21; Mismatches 41; Indels 30; Gaps 6;

```

QY      32 TFSOSPLQKLFARRKKILGTIQLFGIMTSEFVGIFLFTLLKPYRPFPIFLSGYPFWG  91
        | | : : : : | | : | : | : | : |
Db       178 TIIVPEMLKLKKLRVKIRKGHDIK-GIIIMSVGIFF-----MELTTYSIT--  223
                                     TLGILITLMTFSIIELPISL-PSRI 135
        | | : : : : | | : | : | : | : |
Dd       224 SFLIVSYVSLFIWFVHKIKRVDPEDVPGLGNKIILFMIGVLGCGLIFGVAVAGVSVMVPM  283
QY      136 LGCH 139
        |
Db       284 KDVI 287

```

RESULT 11	TCR_STRPN	STANDARD;	PRT;	458 AA.
AC	PI1063; F72219;			
DT	01-JUL-1989 (Rel. 11, Created)			
DT	01-JUL-1989 (Rel. 11, Last sequence update)			
DT	15-DEC-1998 (Rel. 37, Last annotation update)			
DE	Tetracycline resistance protein.			
GN	Ter.			
OS	Streptococcus pneumoniae,			
OS	Bacillus cereus, and			
OS	Bacillus subtilis.			
OG	Plasmid pLS1, Plasmid pBC16, Plasmid pHY163PLK, Plasmid pT919, and			
OC	Plasmid pNS1981			
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;			
OC	Streptococcus.			
OX	NCBI_TaxID=1313, 1396, 1423;			
RM	[1]			
RP	SEQUENCE FROM N.A.			
RC	SPECIES=S.pneumoniae; PLASMID=pLS1;			
RC	MEDLINE=87226167; Pubmed=2438417;			
RA	Lacks S.A., Lopez P., Greenberg B., Espinosa M.;			
RT	"Identification and analysis of genes for tetracycline resistance and			
RT	replication functions in the broad-host-range plasmid pLS1.";			
RL	J. Mol. Biol. 192:753-765(1986).			
RM	[2]			
RP	SEQUENCE FROM N.A.			
RC	SPECIES=B.cereus; PLASMID=pBC16;			
RC	MEDLINE=90221899; Pubmed=2109312;			
RA	Palva A., Vidgren G., Simonen M., Rintala H., Laamanen P.;			
RT	"Nucleotide sequence of the tetracycline resistance gene of pBC16			
RT	from Bacillus cereus.";			
RL	Nucleic Acids Res. 18:1635-1635(1990).			
RM	[3]			
RP	SEQUENCE FROM N.A.			
RC	PLASMID=pHY163PLK;			
RA	Ishida H., Shibahara H.;			
RT	"New shuttle vectors for Escherichia coli and Bacillus subtilis. III.			
RT	Nucleotide sequence analysis of tetracycline resistance gene of			
RT	pma1phal and ori-177.";			
RL	Jpn. J. Genet. 60:485-498(1985).			
RM	[4]			
RP	SEQUENCE FROM N.A.			
RC	PLASMID=pHY163PLK;			
RC	MEDLINE=83129391; Pubmed=6186390;			
RA	Selzer G., Som T., Itoh T., Tomizawa J.;			

"The origin of replication of plasmid p15A and comparative studies on the nucleotide sequences around the origin of related plasmids."; Cell 32:119-129(1983).

[5]

SEQUENCE FROM N.A.

SPECIES=B.subtilis; PLASMID=PTB19;
MEDLINE=92052681; Pubmed-1946749;
Osaka L., Hillenga D.J., Venema G., Bron S.;

"The large Bacillus plasmid pTB19 contains two integrated rolling-circle plasmids carrying mobilization functions."; Plasmid 26:30-39(1991).

[6]

SEQUENCE FROM N.A.

SPECIES=B.subtilis; STRAIN=168 / Marburg; PLASMID=pNS1981;
MEDLINE=86287699; Pubmed-3090576;
Sakauchi R., Shishido K., Hoshino T., Furukawa K.;

"The nucleotide sequence of the tetracycline resistance gene of plasmid pNS1981 from Bacillus subtilis differs from pTET15 from a thermophilic Bacillus by two base pairs."; Plasmid 16:72-73(1986).

-1- FUNCTION: RESISTANCE TO TETRACYCLINE BY AN ACTIVE TETRACYCLINE EFFLUX. THIS IS AN ENERGY-DEPENDENT PROCESS THAT DECREASES THE ACCUMULATION OF THE ANTIBIOTIC IN WHOLE CELLS. THIS PROTEIN FUNCTIONS AS A METAL-TETRACYCLINE/H+ ANTI-PORTER.

-1- SUBCELLULAR LOCATION: Integral membrane protein.

-1- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN AS THE DRUG RESISTANCE TRANSLOCASE FAMILY).

THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/annouce/> or send an email to license@isb-sib.ch).

RT		"The origin of replication of plasmid p15A and comparative studies on
RL		the nucleotide sequences around the origin of related plasmids.";
RN		Cell 32:119-129(1983).
RP	[5]	
RC		SEQUENCE FROM N.A.
RX		SPECIES=B. subtilis; PLASMID=PTB19;
RA		MEDLINE=92052681; PubMed=1946749;
RT		Oskam L., Hillegra D.J., Venema G., Bron S.:
RL		"The large Bacillus plasmid PTB19 contains two integrated
RN		rolling-circle plasmids carrying mobilization functions";
RP		Plasmid 26:30-39(1991).
RC	[6]	
RX		SEQUENCE FROM N.A.
RA		SPECIES=B. subtilis; STRAIN=168 / Marburg; PLASMID=PNS1981;
RT		MEDLINE=86287699; PubMed=3090576;
RL		Sakaguchi R., Shishido K., Hoshino T., Furukawa K.:
RN		"The nucleotide sequence of the tetracycline resistance gene of
RP		plasmid PNS1981 from Bacillus subtilis differs from pTET15 from a
RC		tetrahymonad-like bacterium by two base pairs.";
RX		Plasmid 16:72-73(1986).
RA	-1-	FUNCTION: RESISTANCE TO TETRACYCLINE BY AN ACTIVE TETRACYCLINE
RT		EFFECTOR. THIS IS AN ENERGY-DEPENDENT PROCESS THAT DECREASES THE
RL		ACCUMULATION OF THE ANTIBIOTIC IN WHOLE CELLS. THIS PROTEIN
RN		FUNCTIONS AS A METAL-TETRACYCLINE/H+ ANTIPORTER.
RP	-1-	SUBCELLULAR LOCATION: Integral membrane protein.
RC		-1- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN
RX		AS THE DRUG RESISTANCE TRANSDUCASE FAMILY).
RA		-----
RT		This SWISS-PROT entry is copyright. It is produced through a collaboration
RL		between the Swiss Institute of Bioinformatics and the EMBL Outstation -
RN		the European Bioinformatics Institute. There are no restrictions on its
RP		use by non-profit institutions as long as its content is in no way
RC		modified and this statement is not removed. Usage by and for commercial
RX		entities requires a license agreement (see http://www.isb-sdb.ch/announce/
RA		or send an email to license@sdb.sdb.ch).
RT		-----
RL		EMBL; X51366; CA35751.1; -;
RN		DR EMBL; M29725; AAA98167.1; -;
RP		DR EMBL; D13792; -; NOT_ANNOTATED_CDS.
RC		DR EMBL; M63891; AAA98304.1; -;
RX		EMBL; D00006; BAA00005.1; -;
RA		PIR; S09234; YTB5U6.
RT		PIR; C25599; YTSOG.
RL		PIR; J01211; J01211.
RN		InterPro: IPR001411; TCR_TetB.
RP		PRINTS: PR01036; TCRTETB.
RC		Antibiotic resistance; Transmembrane; Transport; Symport; Plasmid.
RX	KW	TRANSMEM 12 33 POTENTIAL.
RA	FT	TRANSMEM 81 100 POTENTIAL.
RT	FT	TRANSMEM 111 129 POTENTIAL.
RL	FT	TRANSMEM 140 162 POTENTIAL.
RN	FT	TRANSMEM 165 185 POTENTIAL.
RP	FT	TRANSMEM 201 221 POTENTIAL.
RC	FT	TRANSMEM 223 240 POTENTIAL.
RX	FT	TRANSMEM 256 276 POTENTIAL.
RA	FT	TRANSMEM 297 317 POTENTIAL.
RT	FT	TRANSMEM 324 344 POTENTIAL.
RL	FT	TRANSMEM 346 365 POTENTIAL.
RN	FT	TRANSMEM 432 451 POTENTIAL.
RP	FT	SEQUENCE 458 AA; 50092 MW; 6A70877DD4C2074 CRC64;

Query Match		10.4%	Score 80;	DB 1;	Length 458;
Best Local Similarity		25.8%	Pred. No. 3.8;		
Matches	32;	Conservative	21;	Mismatches	41; Indels 30; Gaps 6;
OY	32	TFSQSPLOKLFARKMKILGTQLIFGIMTSPFGIIEFTLLKKPYRRPRFIFLSGYPPWG	91		
Db	178	TIITVPEPLMKLKEVKRIKGHFDIR-GIIIMSVGIFF-----MLFTSYST--	223		
		:	:		
OY	92	SYLEINSGAFILIAVK--RKTE-----TLGLITLMTFEISIEFIISL-PESI	135		
	: :	--	: :	: :	
Db	224	SFLIVASVLSEFLFEAKHKIRKATVDPEYDPOLGKNIPPMIGVLCGGIIEFTVAAGVSAPYMM	283		

OY 136 LGCH 139
 DB 284 KDVI 287

RESULT 12

ADP1_YEAST STANDARD: PRT: 573 AA.
 AC P38971;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Basic amino-acid permease.
 GN ALP1 OR YNL270C OR NO660.
 OS Saccharomyces cerevisiae (Baker's Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OK NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FL100;
 RX MEDLINE=95028162; PubMed=7941748;
 RA Sychrova H., Chevallerier M.R.;
 RT "APL1, a yeast gene encoding a putative permease for basic amino
 acids";
 RL yeast 10:653-657(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / FT1679;
 RX MEDLINE=96310631; PubMed=8740425;
 RA Sen-Gupta M., Lyck R., Fleig U., Nledenthal R.K., Hegemann J.H.;
 RT "The sequence of a 24,152 bp segment from the left arm of chromosome
 XIV from Saccharomyces cerevisiae between the BNI1 and the POL2
 genes";
 RL Yeast 12:505-514(1996).
 CC -1- FUNCTION: HIGH-AFFINITY PERMEASE FOR BASIC AMINO ACIDS.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE AMINO ACID PERMEASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X74069; CAA52199.1; -;
 DR EMBL: X92494; CAA63228.1; -;
 DR EMBL: Z71546; CAA96177.1; -;
 DR PIR: S44329; S44329.
 DR SGD: S0005214; ALP1.
 DR InterPro: IPR002293; AA/rel_pmease1.
 DR InterPro: IPR004840; AAc_permease.
 DR InterPro: IPR004841; Permease.
 DR InterPro: IPR004762; Yeast_AA_perm.
 DR Pfam: PF00324; aa_permeases; 1.
 DR TIGRfam: TIGR00913; 2A0310; 1.
 DR PROSITE: PS00218; AMINO_ACID_PERMEASE_1; 1.
 KW Transport; Amino-acid transport; Transmembrane.
 FT TRANSMEM 75 95
 FT TRANSMEM 96 116
 FT TRANSMEM 154 174
 FT TRANSMEM 180 200
 FT TRANSMEM 210 230
 FT TRANSMEM 266 286
 FT TRANSMEM 307 327
 FT TRANSMEM 360 380
 FT TRANSMEM 404 424
 FT TRANSMEM 432 452
 FT TRANSMEM 478 498
 FT TRANSMEM 511 531
 FT TRANSMEM 511 531
 FT CONFLICT 51 51
 D -> DD (IN REF. 1).

FT CONFLICT 126 126 V -> A (IN REF. 1).
 FT CONFLICT 260 260 D -> N (IN REF. 1).
 FT CONFLICT 517 517 I -> V (IN REF. 1).
 FT CONFLICT 548 548 R -> H (IN REF. 1).
 SQ SEQUENCE 573 AA; 64013 MW; 359DFE1466C348A1 CRC64;

Query Match 10.2%; Score 78.5; DB 1; Length 573;

Best Local Similarity 24.1%; Pred. No. 6.4;

Matches 28; Conservative 24; Mismatches 45; Indels 19; Gaps 4;

OY 20 ASFEYESTELSATFTFQSPLOKLFARKKILGTLQILFEIMFSGVIFLFLNPKYRF 79

DB 277 AFTYQGTGLVGTITAGAPNRRALPRAIK-----KVVRIILVETLISFTFGLVPEY-ND 330

OY 80 PFTLSGYGFQWGSVLEPINSGAFLIAVKRTTEFLG-----LITLTFPSITLLEFL 129

DB 331 PKLDSDG-----ITVSSPFWISLNSGTRVLPDIFNAVVLITLSAGNSNVIT 379

RESULT 13

ANSP_SALTY

ID ANSP_SALTY STANDARD: PRT: 497 AA.

AC P40812;

DT 01-FEB-1995 (Rel. 31, Created)

DT 01-FEB-1995 (Rel. 31, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE L-asparagine permease (L-asparagine transport protein).

GN ANSP OR STM1584.

OS Salmonella typhimurium.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Salmonella.

OK NCBI_TaxID=602;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=SA2656;

RX MEDLINE=95202072; PubMed=7894705;

RA Jennings P.J., Anderson J.K., Beacham I.R.;

RT "Cloning and molecular analysis of the Salmonella enterica ansp gene,
 encoding an L-asparagine permease";

RL Microbiology 141:141-146(1995).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=LT2 / SGSC1412 / ATCC 700720;

RX MEDLINE=21534948; PubMed=11677609;

RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,

RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,

RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,

RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,

RA Waterston R., Wilson R.K.;

RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
 LT2";

RL Nature 413:852-856(2001).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.

CC -1- SIMILARITY: BELONGS TO THE AMINO ACID PERMEASE FAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: U04851; AAA80001.1; -;

DR EMBL: AE008769; AAL20502.1; -;

DR StyGene: SG10527; ansp.

DR InterPro: IPR002293; AA/rel_pmease1.

DR InterPro: IPR004840; AAc_permease.

DR InterPro: IPR004841; Permease.

DR Pfam: PF00324; aa_permeases; 1.

DR PROSITE: PS00218; AMINO_ACID_PERMEASE_1; 1.

DR TRANSPORT; Amino-acid transport; Transmembrane; Inner membrane;

FT TRANSMEM 51 51

FT TRANSMEM 51 51

FT TRANSMEM 51 51

FT TRANSMEM 51 51

FT TRANSMEM 51 51

FT TRANSMEM 51 51

FT TRANSMEM 51 51

FT TRANSMEM 51 51

FT TRANSMEM 51 51

FT TRANSMEM 51 51

FT TRANSMEM 51 51

FT TRANSMEM 51 51

FT TRANSMEM 51 51

FT TRANSMEM 34 54 POTENTIAL.
FT TRANSMEM 58 78 POTENTIAL.
FT TRANSMEM 109 129 POTENTIAL.
FT TRANSMEM 146 166 POTENTIAL.
FT TRANSMEM 171 191 POTENTIAL.
FT TRANSMEM 219 239 POTENTIAL.
FT TRANSMEM 264 284 POTENTIAL.
FT TRANSMEM 298 318 POTENTIAL.
FT TRANSMEM 353 373 POTENTIAL.
FT TRANSMEM 378 398 POTENTIAL.
FT TRANSMEM 422 442 POTENTIAL.
FT TRANSMEM 448 468 POTENTIAL.
SQ SEQUENCE 497 AA; 54004 MW; BE80CB117992CBB6 CRC64;

Query Match 10.1%; Score 77.5; DB 1; Length 497;
Best Local Similarity 22.8%; Pred. No. 6.9; Mismatches 25; Indels 53; Gaps 8;

QY 23 YESTELSATFTSTQSPLOKLFARKMKILGTIOILFGIMTFSFGVILFTLLKPPRP--- 79
DB 233 FASIEIVGTAGACBCKDPKQKVPKAIN-----SVIWRIGLFYGVSVLLVLLPMNNAVQAG 287
QY 80 --PFI-FLS--GYPFNGSVLFT-----NSG-----AFLIAVK 106
DB 288 QSPFVTFEFSKLVGYPIGSIINIVLTAALSLNSGLYCTGRILRSMSGGAPKFAKMS 347
QY 107 RKTEETLILITLMTFSIIEFLFIS-----LPFSILG 137
DB 348 ROHVPYAGILATLVY-VGVFLNLYVPSRVEITVLNFSLSG 388

RESULT 14

COTE HUMAN

ID COTE HUMAN STANDARD; PRT; 669 AA.

AC P81408; 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE COTE1 protein.
GN CLORE2 OR COTE1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hippocampus;
RA Winfield S.L., Tayebi N., Martin B.M., Gims E.I., Sidransky E.;
RT "Identification of three additional genes contiguous to the
RT glucocerebrosidase locus on chromosome 1q21: implications for Gaucher
RT disease.";
RL Genome Res. 7:1020-1026(1997).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED.
CC -1- SIMILARITY: TO HUMAN KIA00574.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF023268; AAC51822.1; -
DR Genew; HGNC:1233; CLOF2.
KW Transmembrane.
FT TRANSMEM 34 54 POTENTIAL.
FT TRANSMEM 67 87 POTENTIAL.
FT TRANSMEM 91 111 POTENTIAL.
FT TRANSMEM 174 194 POTENTIAL.
FT DOMAIN 244 250 POLY-PRO.

FT DOMAIN 635 639 POLY-SER.
SQ SEQUENCE 669 AA; 71482 MW; E30360AC9A4571E6 CRC64;
Query Match 10.0%; Score 77; DB 1; Length 669;
Best Local Similarity 22.6%; Pred. No. 10;
Matches 31; Conservative 25; Mismatches 67; Indels 14; Gaps 5;

QY 16 PETASVESTELSATFTSTQSPLOKLFARKMKILGTIOILFGIMTFSFGVILFTLLK 74
DB 3 PPSDSRSRLTSPSTIGTLHLRHWLQALLTLGLVGLLIVYTESMVASVTTTE 62
QY 75 PYPFPFIFLUSGYPFNGSVLFINSGAF-LIAYRKTEETLILITLMTFSIIEFLSLPP 133
DB 63 SIKR-----SCPSMAGFSLAFSGVGYVMKRPFTLVISFSL---SVLCVMSMAG 112
QY 134 SITGCHSEDC--DCBEC 148
DB 113 SVLSCKNAQLARDFOOC 129

RESULT 15

NUSM_ARATH

ID NUSM_ARATH STANDARD; PRT; 669 AA.

AC P29388; 01-DEC-1992 (Rel. 24, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 5 (EC 1.6.5.3).
GN ND5 OR NAD5.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Unseled M., Martienfeld J.R., Brandt P., Brennicke A.;
RT "The mitochondrial genome of Arabidopsis thaliana contains 57 genes
RT in 366,924 nucleotides.";
RL Nucleic Acids Res. 25:57-61(1997).
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X60045; CAA42648.1; -
DR EMBL; X60047; CAA42648.1; JOINED.
DR EMBL; X60048; CAA42648.1; JOINED.
DR EMBL; Y08501; CAA69752.3; -
DR PIR; S20234; DNMU05.
DR InterPro: IPR003916; NADH-ub.oxred5.
DR InterPro: IPR001750; Oxidored.q1.
DR InterPro: IPR001516; Oxidored.q1_N.
DR Pfam; PF00361; oxidored.q1; 1.
DR Pfam; PF00662; oxidored.q1_N; 1.
DR PRINTS; PR01434; NADHGNAS5.
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.

SQ SEQUENCE 669 AA; 73907 MN; 0748DCD81DB40429 CRC64;

```
Query Match      9.8%; Score 75.5; DB 1; Length 669;
Best local similarity 34.3%; Prod No 14.
```

Best Local Similarity 24.3%; Pred. No. 14;

Matches 36; Conservative 30; Mismatches 51; Indels 31; Gaps 8;

8 SPVFLVFPPE---ITASEYESTELSATTFSTQSPLOKLFA-RKKKILGTIQILLFGIMTF 62

Db 273 SPLF-EYPTALIVTSAGATTSLAATTGILQNDLKRVIAYSTCSQLGYMIFACGISNY 331

QY 63 SEGVIFLETLTKPRFPFIELSGYPFWGSVLFINSGLIAVK-----RK---TTEELG 114

Db 332 SVSVFHLM-----NHAFKALLFLSAGSVIHAMSDEQDMRKMGGLASSFP 376

115 ILITMTFSIIEFLSPSILGCHSED 142Z

Db 377 LTYAMMLIGSLSL-IGPF-LTGFSKD 402

```
Search completed: February 24, 2003, 13:03:35
Job time : 16.1293 secs
```

Job time : 16.1293 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2003, 13:00:36 ; Search time 23.1207 Seconds
(without alignments)
1327.859 Million cell updates/sec

Title: US-09-735-712-8

Perfect score: 771

Sequence: 1 MDSTASHPVFLVFPPEITTA.....SLPFLIGCHSEDCDEQCC 149

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database :

SPREMBL_21:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp rodent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*
15: sp.virus:*
16: sp.bacteriaph:*
17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	735.5	95.4	200	4 Q9B2H1	Q9B2H1 homo sapien
2	735.5	95.4	200	4 Q9H3V2	Q9H3V2 homo sapien
3	161.5	20.9	167	4 Q96PG7	Q96PG7 homo sapien
4	155	20.1	679	4 Q96J44	Q96J44 homo sapien
5	154.5	20.0	214	4 Q96HJ5	Q96HJ5 homo sapien
6	145.5	18.9	217	11 Q9D729	Q9D729 mus musculu
7	143.5	18.6	217	11 Q99N08	Q99N08 mus musculu
8	141.5	18.4	178	4 Q9H3V1	Q9H3V1 homo sapien
9	139	18.0	225	4 Q9HC76	Q9HC76 homo sapien
10	139	18.0	225	4 Q96PG6	Q96PG6 homo sapien
11	139	18.0	248	4 Q9H2M1	Q9H2M1 homo sapien
12	139	18.0	248	4 Q8TBV7	Q8TBV7 homo sapien
13	138	17.9	225	4 Q8TE24	Q8TE24 homo sapien
14	138	17.9	248	4 Q8TE25	Q8TE25 homo sapien
15	137.5	17.8	226	11 Q9ES61	Q9ES61 mus musculu
16	134.5	17.4	268	11 Q99N10	Q99N10 mus musculu

17	134	17.4	197	4 Q8TE26	Q8TE26 homo sapien
18	134	17.4	205	4 Q9H3V3	Q9H3V3 homo sapien
19	134	17.4	213	11 Q920C4	Q920C4 mus musculu
20	134	17.4	220	4 Q9B1T8	Q9B1T8 homo sapien
21	134	17.4	239	4 Q96J05	Q96J05 homo sapien
22	133.5	17.3	225	11 Q99N05	Q99N05 mus musculu
23	133	17.3	135	6 Q95L74	Q95L74 macaca fasc
24	132	17.1	197	4 Q9PIS3	Q9PIS3 homo sapien
25	131.5	17.1	244	11 Q99N09	Q99N09 mus musculu
26	131.5	17.1	249	4 Q9H2N3	Q9H2N3 homo sapien
27	131	17.0	220	11 Q9D8E5	Q9D8E5 mus musculu
28	130.5	16.9	124	11 Q99MX6	Q99MX6 mus musculu
29	129.5	16.8	226	11 Q9D3F6	Q9D3F6 mus musculu
30	129.5	16.8	244	11 Q9B8W9	Q9B8W9 mus musculu
31	128	16.6	267	4 Q9KXJ0	Q9KXJ0 homo sapien
32	125.5	16.3	268	11 Q9E0Z0	Q9E0Z0 mus musculu
33	122.5	15.9	190	11 Q9E0Y7	Q9E0Y7 mus musculu
34	122.5	15.9	230	11 Q9E0Y9	Q9E0Y9 mus musculu
35	122.5	15.9	247	11 Q99N07	Q99N07 mus musculu
36	119.5	15.5	116	11 Q8R046	Q8R046 mus musculu
37	119.5	15.5	247	11 Q9D0X1	Q9D0X1 mus musculu
38	116.5	15.1	125	11 Q91ZV5	Q91ZV5 mus musculu
39	115.5	15.0	250	4 Q8TCAS	Q8TCAS homo sapien
40	114.5	14.9	240	4 Q9BX19	Q9BX19 homo sapien
41	105.5	13.7	240	4 Q9GZM8	Q9GZM8 homo sapien
42	103.5	13.4	249	4 Q9H2L1	Q9H2L1 homo sapien
43	100.5	13.0	176	11 Q9DC76	Q9DC76 mus musculu
44	100	13.0	457	5 Q76620	Q76620 caenorhabd
45	95.5	12.4	234	11 Q99N04	Q99N04 mus musculu

ALIGNMENTS

RESULT 1
Q9B2H1 PRELIMINARY: PRT: 200 AA.
AC Q9B2H1;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE Testis-expressed transmembrane-4 protein.
GN TEM4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-21092614; PubMed-11162526;
RA Hulet M.D., Pagler E., Hornby J.R., Hogarth P.M., Eyre H.J.,
BA Baker E., Crawford J., Sutherland G.R., Ohms S.J., Parish C.R.;
RT "Isolation, tissue distribution, and chromosomal localization of a
RT novel testis-specific human four-transmembrane gene related to CD20
RT and Fcgpst10nrt-beta.";
RT Biochem. Biophys. Res. Commun. 280:374-379(2001).
DR EMBL: AF321127; AAK01641.1; -;
KW Transmembrane.
SQ SEQUENCE 200 AA; 22249 MW; 7B282E5D15B25BF4 CRC64;
QY
Query Match 95.4%; Score 735.5; DB 4; Length 200;
Best Local Similarity 74.5%; Pred. No. 6.3e-67;
Matches 149; Conservative 0; Mismatches 0; Indels 51; Gaps 1;
QY 1 MDSTASHPVFLVFPPEITASEYESTELSATPSTOSPLQKLPARKMKILGTIOILFGIM 60
|||||
1 MDSTASHPVFLVFPPEITASEYESTELSATPSTOSPLQKLPARKMKILGTIOILFGIM 60
QY 61 TFSFGVIFLTLKPYPRPFIFLSCYPWGSVLFINSGAFILAVRKRTTET----- 112
|||||
61 TFSFGVIFLTLKPYPRPFIFLSCYPWGSVLFINSGAFILAVRKRTTETLILISRIIM 120
|||||
QY 113 -----LGIILITMPSITIELFI 129

RM [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC008487; AAOH8487.1; -
SQ SEQUENCE 214 AA; 22933 MW; D30C276213DB8F3D CRC64;

Query Match 20.0%; Score 154.5; DB 4; Length 214;
Best Local Similarity 31.7%; Pred. No. 6.1e-08;
Matches 52; Conservative 23; Mismatches 62; Indels 27; Gaps 8;

QY 1 MDSSTAH-SPEFLVPEPEITASEESTELSTSTFSPLOKLFARKMKILGTIQLFI 59
ID 099N08 PRELIMINARY; PRT; 217 AA.
AC 099N08;
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DE 60 MTFSEGVIFLETLTKPY----PRPFITLGGYPPWGSVLFINSGAFILAVKRKTE--- 112
DB 61 MIALGV-FLGSLQYPIHFKHFFFTFYTGYPIMGAVFCCSGTSLVAGIKPFTWIQ 119
PY 113 --LGILITMTFSIT-ELFISLPF----SLGCHSEDCDCEC 148
DB 120 NSFQMNIAATIALVGTAFILSLNIAVNIQSILRSCHSSESPDL 163

RESULT 6
Q99DZ9
ID 099DZ9 PRELIMINARY; PRT; 217 AA.
AC 099DZ9;
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DE 01-JUN-2001 (TEMBLrel. 17, last sequence update)
DE 2200009H22RIK protein.
GN 2200009H22RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIB-C57BL/6J; TISSUE=STOMACH;
RX MEDLINE=21085650; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Akawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nakado I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Mashio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Wetz C., Whitaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AK008652; BAH25808.1; -
DR MGD; MGI:1920906; 2200009H22RIK
SQ SEQUENCE 217 AA; 23643 MW; 2E6C4FF0287B543F CRC64;

Query Match 18.9%; Score 145.5; DB 11; Length 217;
Best Local Similarity 28.4%; Pred. No. 5e-07;
Matches 44; Conservative 31; Mismatches 59; Indels 21; Gaps 5;

QY 2 DSSTAHSPVLPVPPETASEESTELSATTFSTQSPLOKLFARKMKILGTIQLFI 61
DB 9 ETTITSPNGINFPQK-----DESQPTQQRDSILKRLKAEIVIAIQIMCAVTV 59

QY 62 FSEGVIFLETLTKPYPPPFIF-----LSGYPPWGSVLFINSGAFILAVKRKTE-----T 112
DB 60 LALGIIILASG--PPVFNFSVFLSKSGYFICALFFIASGIIITTEKSTKLVDAS 117
QY 113 LGILITMTFSITIELFISLPFISLIGCHSEDCDCEQ 147
DB 118 LTLNITLSVFAFVGII-IVSLVGLHPASBQCKQ 151

RESULT 7
Q99N08
ID 099N08 PRELIMINARY; PRT; 217 AA.
AC 099N08;
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DE 01-JUN-2001 (TEMBLrel. 17, last sequence update)
DE 01-DEC-2001 (TEMBLrel. 19, last annotation update)
DE MS4AG protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TOTAL FETUS;
RX MEDLINE=21295030; PubMed=11401424;
RA Liang Y., Tedder T.F.;
RT "Identification of a cd20-, fcepsilon-related, and htm4-related gene
family: sixteen new ms4a family members expressed in human and
mouse.";
RL Genomics 72:119-127(2001).
DR EMBL: AF237910; AAK37419.1; -
SQ SEQUENCE 217 AA; 23622 MW; 8BF521AF22DBB7BD CRC64;

Query Match 18.6%; Score 143.5; DB 11; Length 217;
Best Local Similarity 27.7%; Pred. No. 8.1e-07;
Matches 43; Conservative 31; Mismatches 60; Indels 21; Gaps 5;

QY 2 DSSTAHSPVLPVPPETASEESTELSATTFSTQSPLOKLFARKMKILGTIQLFI 61
DB 9 ETTITSPNGINFPQK-----DESQPTQQRDSILKRLKAEIVIAIQIMCAVTV 59

QY 62 FSEGVIFLETLTKPYPPPFIF-----LSGYPPWGSVLFINSGAFILAVKRKTE-----T 112
DB 60 LALGII-LSAVPPVFNFSVFLSKSGYFICALFFIASGIIITTEKSTPLVDAS 117
QY 113 LGILITMTFSITIELFISLPFISLIGCHSEDCDCEQ 147
DB 118 LTLNITLSVFAFVGII-IVSLVGLHPASBQCKQ 151

RESULT 8
Q99H3V1
ID 099H3V1 PRELIMINARY; PRT; 178 AA.
AC 099H3V1;
DT 01-MAR-2001 (TEMBLrel. 16, Created)
DE 01-MAR-2001 (TEMBLrel. 16, last sequence update)
DE 01-DEC-2001 (TEMBLrel. 19, last annotation update)
DE MS4AG.
GN MS4AG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21142397; PubMed=11245982;
RA Ishibashi K., Suzuki M., Sasaki S., Imai M.;
RT "Identification of a new multigene four-transmembrane family (MS4A)
related to CD20, Htm4 and beta subunit of the high-affinity IgE
receptor.";
RL Gene 264:87-93(2001).
DR EMBL: AB013104; BAB18740.1; -

SQ SEQUENCE 178 AA; 18801 MW; 54E10028530D914 CRC64;
 Query Match 18.4%; Score 141.5; DB 4; Length 178;
 Best Local Similarity 26.2%; Pred. No. 1.1e-06;
 Matches 45; Conservative 28; Mismatches 74; Indels 25; Gaps 6;
 QY 1 MDSSTAHSPVFLVPPETITASEYESTELSATFTSTQSPLOKLFARKMKILGTIOILFGIM 60
 DB 1 MTSQPVNPTIIVLPNSVI--NFSQAEKPEPTNOGDSLKHHAEIKVIGTIOILCGMM 58
 QY 61 TFSFGVIFLTLKP-YRPPFTFL-SGYPFGSVLFINSAGFLIAVKRKTTETL----- 113
 DB 59 VLSGLITLASASFSPTQVSTLNSAVPFIQFPFFIISGSLSIATEKRLTKLVHSSL 118
 QY 114 -GILITLMTFSIIEFLFISL-----PFSILGCHSEDCD-----CEOC 149
 DB 119 VGSILSALSALVGFIILSVKQATLNPAISLQMSLSDADLHSAITLPCACMC 170
 RESULT 9
 ID 09HC76 PRELIMINARY; PRT; 225 AA.
 AC 09HC76;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)
 DE CD20-1-like precursor.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NC NCB1;TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Zhang W., Wan T., He L., Yuan Z., Cao X.;
 RT "Novel human CD20-1-like molecule."
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF142409; AAG27920.1; -
 SQ SEQUENCE 225 AA; 24317 MW; 144D2EB7DC71B56A CRC64;
 Query Match 18.0%; Score 139; DB 4; Length 225;
 Best Local Similarity 26.5%; Pred. No. 2.4e-06;
 Matches 41; Conservative 31; Mismatches 73; Indels 10; Gaps 4;
 QY 1 MDSSTAHSPVFLVPPETITASEYESTELSATFTSTQSPLOKLFARKMKILGTIOILFGIM 60
 DB 1 MTSQPVNPTIIVLPNSVI--NFSQAEKPEPTNOGDSLKHHAEIKVIGTIOILCGMM 58
 QY 61 TFSFGVIFLTLKP-YRPPFTFL-SGYPFGSVLFINSAGFLIAVKRKTTETL----- 113
 DB 59 VLSGLITLASASFSPTQVSTLNSAVPFIQFPFFIISGSLSIATEKRLTKLVHSSL 118
 QY 114 -GILITLMTFSIIEFLFISL-----PFSILGCHSEDCDCEQ 147
 DB 119 VGSILSALSALVGFIILSVKQATLNPAISLQCELDK 153
 RESULT 10
 ID 096PG6 PRELIMINARY; PRT; 225 AA.
 AC 096PG6;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
 DE MS4A6A.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NC NCB1;TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RX MEDLINE=21382044; PubMed=11486273;
 RX MEDLINE=21382044; PubMed=11486273;
 RA Liang Y., Buckley T.R., Tu L., Langdon S.D., Tedder T.F.;

RT "Structural organization of the human MS4A gene cluster on Chromosome
 11q12."
 RL Immunogenetics 53:357-368(2001).
 DR EMBL: AF354930; AAL07357.1; -
 SQ SEQUENCE 225 AA; 24331 MW; 12FBF5DDC71B56A CRC64;
 Query Match 18.0%; Score 139; DB 4; Length 225;
 Best Local Similarity 26.5%; Pred. No. 2.4e-06;
 Matches 41; Conservative 31; Mismatches 73; Indels 10; Gaps 4;
 QY 1 MDSSTAHSPVFLVPPETITASEYESTELSATFTSTQSPLOKLFARKMKILGTIOILFGIM 60
 DB 1 MTSQPVNPTIIVLPNSVI--NFSQAEKPEPTNOGDSLKHHAEIKVIGTIOILCGMM 58
 QY 61 TFSFGVIFLTLKP-YRPPFTFL-SGYPFGSVLFINSAGFLIAVKRKTTETL----- 113
 DB 59 VLSGLITLASASFSPTQVSTLNSAVPFIQFPFFIISGSLSIATEKRLTKLVHSSL 118
 QY 114 -GILITLMTFSIIEFLFISL-----PFSILGCHSEDCDCEQ 147
 DB 119 VGSILSALSALVGFIILSVKQATLNPAISLQCELDK 153
 RESULT 11
 ID 09H2W1 PRELIMINARY; PRT; 248 AA.
 AC 09H2W1;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)
 DE CD401 (MS4A6A-POLYMORPH) (MS4A6A protein).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NC NCB1;TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=PHEOCHROMOCYTOMA;
 RA Li Y., Huang Q., Peng Y., Song H., Yu Y., Xu S., Ren S., Chen Z.,
 Han Z.;
 RT Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 RN (2)
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER, AND SPLEEN;
 RX MEDLINE=21295030; PubMed=11401424;
 RA Liang Y., Tedder T.F.;
 RT "Identification of a CD20-, FcγRIIIa-related, and HTM4-related Gene
 Family: Sixteen New MS4A Family Members Expressed in Human and
 Mouse."
 RL Genomics 72:119-127(2001).
 DR EMBL: AF212240; AAG41780.1; -
 DR EMBL: AF286866; AAK37994.1; -
 DR EMBL: AF237908; AAK37417.1; -
 SQ SEQUENCE 248 AA; 26943 MW; 49892BDC25D20A23 CRC64;
 Query Match 18.0%; Score 139; DB 4; Length 248;
 Best Local Similarity 26.5%; Pred. No. 2.6e-06;
 Matches 41; Conservative 31; Mismatches 73; Indels 10; Gaps 4;
 QY 1 MDSSTAHSPVFLVPPETITASEYESTELSATFTSTQSPLOKLFARKMKILGTIOILFGIM 60
 DB 1 MTSQPVNPTIIVLPNSVI--NFSQAEKPEPTNOGDSLKHHAEIKVIGTIOILCGMM 58
 QY 61 TFSFGVIFLTLKP-YRPPFTFL-SGYPFGSVLFINSAGFLIAVKRKTTETL----- 113
 DB 59 VLSGLITLASASFSPTQVSTLNSAVPFIQFPFFIISGSLSIATEKRLTKLVHSSL 118
 QY 114 -GILITLMTFSIIEFLFISL-----PFSILGCHSEDCDCEQ 147
 DB 119 VGSILSALSALVGFIILSVKQATLNPAISLQCELDK 153
 RESULT 12

